Infinite Sets

Lecture notes for NMAI074

Jakub Smolík

smolikj@matfyz.cz

Date of last change: November 16, 2025

Contents

1 Review of Set Theory Basics **2**
- 1.1 Sets and Classes . . . 2
- 1.2 Axiom Schema of Replacement . . . 2
- 1.3 Axiom of Choice . . . 3
- 1.4 Natural Numbers and the Axiom of Infinity . . . 5
- 1.5 Well-Orderings and Initial Segments . . . 5

2 Ordinal Numbers **7**
- 2.1 Ordinals as a Generalization of Naturals . . . 7
- 2.2 Ordinals as Types of Well-Ordered Sets . . . 8
- 2.3 Transfinite Induction and Recursion . . . 10
- 2.4 The Well-Ordering Principle . . . 12
- 2.5 Zorn's Lemma . . . 13
- 2.6 The Trichotomy Principle . . . 16

3 Operations on Ordinals **17**
- 3.1 Ordinal Functions . . . 17
- 3.2 Ordinal Arithmetic . . . 20
  - 3.2.1 Definitions and Intuition . . . 20
  - 3.2.2 Basic Properties of Ordinal Operations . . . 22
  - 3.2.3 Ordinal Equations and Power Expansions . . . 26
- 3.3 Countable and Uncountable Ordinals . . . 28
  - 3.3.1 Epsilon Numbers . . . 28
  - 3.3.2 The Veblen Hierarchy . . . 31
  - 3.3.3 Hartogs' Theorem . . . 33
- 3.4 Peano Arithmetic . . . 35
  - 3.4.1 Peano Axioms . . . 35
  - 3.4.2 Models of Arithmetic . . . 36
  - 3.4.3 Gödel's Incompleteness Theorems . . . 37
  - 3.4.4 Consistency and the Connection with ε_0 . . . 39
  - 3.4.5 Limits of Predicative Mathematics . . . 40
- 3.5 Applications of Countable Ordinals . . . 41
  - 3.5.1 Goodstein Sequences . . . 41
  - 3.5.2 The Hydra Game . . . 46
  - 3.5.3 Fundamental Sequences . . . 47
  - 3.5.4 The Fast-Growing Hierarchy . . . 48

4 Cardinal Numbers **52**
- 4.1 Cardinals as Sizes of Well-Ordered Sets . . . 52
- 4.2 Infinite Cardinals . . . 54
- 4.3 Cofinality and Inaccessible Cardinals . . . 58
- 4.4 The Continuum Hypothesis . . . 61
- 4.5 Cardinal Arithmetic . . . 61

Sources **62**

1 Review of Set Theory Basics

We will be working within Zermelo–Fraenkel (ZF) set theory; that is, Zermelo's (1908) theory Z[1] augmented [2] by Fraenkel's (1922) axiom schema of replacement and von Neumann's (1925) axiom of foundation. If we further include the axiom of choice, we obtain the much stronger theory ZFC.

As for notation, I always use the symbol $\subset$ for a proper subset (or subclass) and $\subseteq$ for a general subset (or subclass). I use $\subsetneq$ only when it is important that the two sets (or classes) are not equal. Concatenated expressions such as $a \in b \in c$ mean $a \in b \wedge b \in c$. I differentiate between the symbol for equality of two objects "=" and the symbol for the definition of an object ":=". I use the following notation for defining functions.

- $f : A \to B$ is a function with domain A and codomain B.
- $f : a \mapsto b$ denotes that f maps the set $a \in A$ to the set $b \in B$.
- $f = g \circ h$ means that $f(x) = h(g(x))$ for all suitable x.

I use the terms "function," "map," and "mapping" interchangeably.

1.1 Sets and Classes

Definition 1.1 (Class)**.** If $\varphi(x)$ is a formula, then the expression $\{x \,|\, \varphi(x)\}$ is called a *class term*. It defines the "collection" of all sets x satisfying $\varphi(x)$. We call this collection the *class* determined by $\varphi(x)$.

Every set is a class, but not all classes are sets (consider the class of all sets). A class that is not a set is called a *proper class*. The major difference between sets and classes is that classes cannot be members of other classes or sets, while sets can. We can substitute class terms into logical formulas in place of free variables, but unlike sets, we cannot quantify them using $\forall$ and $\exists$. It isn't hard to show that for every formula with class terms (but without quantified class variables), there is an equivalent formula in the base language without class terms.

We will usually denote sets using small letters $a, b, c, x, y, \ldots$ and classes using capital letters A, B, C, etc. The exception to this are well-ordered sets, which will often be denoted as W. Finally, the class of all sets, also called the *universal class*, is denoted by V.

1.2 Axiom Schema of Replacement

Axiom 1.2. When we take any (even a class) map F and a preimage set a, then the class of images $b = F[a]$ is also a set. Formally, if $\psi(x,y)$ is a formula without free variables y_1, y_2 and b, then

$$(\forall x)(\forall y_1,y_2)\big((\psi(x,y_1) \wedge \psi(x,y_2)) \Rightarrow y_1 = y_2\big) \Rightarrow \\ (\forall a)(\exists b) : (\forall y)\big(y \in b \Leftrightarrow (\exists x)(x \in a \wedge \psi(x,y))\big)$$

[1] Zermelo's theory Z actually already contained AC, which he formulated in 1904.

[2] Replacement is necessary to prove the existence of some key sets, as demonstrated in Section 3.3.3. Foundation is more of a "cleanup" axiom, as virtually all results in the branches of mathematics based on set theory hold even without it.

is an axiom. The formula $\psi(x,y_1)$, resp. $\psi(x,y_2)$ are created from $\psi(x,y)$ by substituting y_1, resp. y_2 for y.

The first part of this axiom says that $\psi(x,y)$ should behave like a map $y = F(x)$. In the second part, a denotes the set of preimages and b the set of corresponding images.

1.3 Axiom of Choice

The *axiom of choice*, denoted AC, is one of the most important principles in modern mathematics, with profound implications in areas such as classical analysis or linear algebra. It states that for any collection of nonempty sets, it is possible to choose exactly one element from each set, even if the collection is infinite. When added to Zermelo–Fraenkel set theory, it yields the much more powerful ZFC. Many theorems that seem intuitively true, such as every vector space having a basis, depend on this axiom.

However, the axiom of choice is also controversial, as it leads to counter-intuitive results, such as the well-ordering principle, which claims that every set can be well-ordered, or the Banach–Tarski paradox, which provides a way to decompose a solid ball into finitely many pieces and reassemble them into two identical copies of the original.

Definition 1.3 (Choice function)**.** A *choice function* (or a *selector*) on the set x is any function $f : x \to \bigcup x$ such that

$$(\forall t \in x)(t \neq \varnothing \Rightarrow f(t) \in t).$$

We can WLOG assume that the choice function is defined on $x \setminus \{\varnothing\}$ and all $t \in \mathrm{Dom}(f)$ satisfy $f(t) \in t$.

One can prove in ZF via finite induction that every finite set has a choice function; that is, we are allowed to make finitely many choices (out of potentially infinite sets). However, it can be shown that ZF cannot prove that every countable set has a choice function, and certainly not that *every* set has a choice function.

Since the assumption that *every* set has a choice function can lead to some paradoxical results (such as the Banach–Tarski paradox), we distinguish three different "power levels" of this axiom.

Axiom 1.4 (Axiom of Countable Choice AC_ω)**.** Every countable set has a choice function; one can make only a countable number of choices.

Axiom 1.5 (Axiom od Dependent Choice DC)**.** One can make a countable sequence of choices, where each choice may *depend* on the previous ones. Formally, for any nonempty set A and a binary relation $R \subseteq A \times A$ such that

$$(\forall x \in A)(\exists y \in A)\, x\, R\, y,$$

there exists an infinite sequence (x_n) satisfying $x_n\, R\, x_{n+1}$ for all $n \in \omega$.

Axiom 1.6 (Axiom of Choice AC)**.** Every set has a choice function; the amount of choices is not limited.

Exercise 1. Show that $\mathsf{AC} \implies \mathsf{DC} \implies \mathsf{AC}_\omega$.

We will now present a brief overview of some of the results that can be proven from each power level (but cannot be proven in ZF). More details and proofs can be found in [11].

Axiom of countable choice AC_ω

$\iff$ an arbitrary Cartesian product of countably many nonempty sets is nonempty.

$\implies$ any union of a countable collection of countable sets is countable.

$\implies$ every infinite set has a countable subset. Or equivalently, there is no infinite set x such that $x \prec \omega$

$\implies$ every set x is finite $\iff$ it is Dedekind finite; that is $(\forall y)(y \subset x \Rightarrow y \prec x)$.

$\implies$ a function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point x $\iff$ for every convergent sequence (x_n) we have $\lim f(x_n) = f(\lim x_n)$.

Axiom of dependent choice DC

$\implies$ a linearly ordered set is well-ordered $\iff$ it contains no infinite strictly descending sequences.

$\implies$ most results of classical analysis and topology.

Axiom of choice AC

$\iff$ an arbitrary Cartesian product of nonempty sets is nonempty.

$\implies$ there exists a mapping which assigns to each set x a set $|x|$ such that $x \approx |x|$ and $x \approx y \iff |x| = |y|$.

$\iff$ for any infinite set x it holds that $|x| = |x \times x|$.

$\iff$ every vector space (even of infinite dimension) has a basis.

$\iff$ the product of (even infinitely many) compact topological spaces is compact.

$\iff$ every surjection $f : X \to Y$ has a right inverse, i.e. a function $g : Y \to X$ such that $f(g(y)) = y$ for all $y \in Y$.

$\iff$ every connected (even infinite) graph has a spanning tree.

$\implies$ the Compactness Theorem in first-order logic: if every finite subset of a theory T has a model, then T has a model.

$\implies$ the uncountable set $\mathbb{R}$ can be well-ordered.

$\iff$ **the Well-Ordering Principle**: every set can be well-ordered.

$\iff$ **Zorn's Lemma**: every ordered set containing upper bounds for every chain necessarily contains at least one maximal element.

$\iff$ **the Trichotomy Principle**: for any sets x and y either $x \preceq y$, or $y \preceq x$.

We will show the equivalence[3] of AC and the last three conditions in Sections 2.4, 2.5 and 2.6.

[3] "The Axiom of Choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?" — Jerry Bona

Paradoxical results implied by AC

$\Longrightarrow$ the Banach–Tarski paradox: it is possible to decompose a solid ball into a few pieces and reassemble them into two identical copies of the original ball. Vsauce has a VIDEO with an intuitive explanation.

$\Longrightarrow$ there exist subsets of $\mathbb{R}$ that are not Lebesgue measurable. The most famous such set is probably the Vitali set. Veritasium has a great VIDEO on this topic (and the history of AC in general).

The axioms of countable and dependent choice are implicitly used in disciplines such as classical analysis all the time. The full power of the axiom of choice is rarely needed, and we try to avoid it when possible.

1.4 Natural Numbers and the Axiom of Infinity

We use Von Neumann ordinals, meaning that natural numbers are defined as

$$0 \coloneqq \varnothing,\ 1 \coloneqq \{0\},\ 2 \coloneqq \{0,1\},\ \dots,\ n+1 \coloneqq \{0,1,\dots,n\} = n \cup \{n\}.$$

Definition 1.7. The *successor function* is a mapping $S : V \to V$ defined as $v \mapsto v \cup \{v\}$. For convenience, we write $v + 1 \coloneqq S(v) = v \cup \{v\}$.

Definition 1.8. A set w is *inductive* if $0 \in w$ and for all $n \in w$ also $n + 1 \in w$.

Axiom 1.9 (Axiom of Infinity)**.** There exists an inductive set.

Definition 1.10. We define the *set of all natural numbers* as the $\subseteq$–smallest inductive set. Or, equivalently, as $\bigcap\{w \mid w \text{ is inductive}\}$. We denote it by ω.

1.5 Well-Orderings and Initial Segments

Let us recall a very important definition: the notion of *well-ordered sets*.

Definition 1.11 (Ordering)**.** A binary relation R on the class X is a

(a) *trichotomy* if for all x,$y \in X$, either $x = y$, or $x\,R\,y$, or $y\,R\,x$,

(b) *strict order* if it is anti-reflexive, strongly anti-symmetric, and transitive on X; (note that strong anti-symmetry follows from the other two),

(c) *(partial) order* if it is reflexive, weakly anti-symmetric, and transitive on X,

(d) *total (or linear) order* if it is a trichotomous partial order on X.

If R is an ordering, then instead of $x\,R\,y$ we write $x \leq_R y$ and we call $(X, \leq_R)$ an *ordered class*. Similarly, if R is a strict ordering, then we write $x <_R y$ and we call $(X, <_R)$ a *strictly ordered class*.

Note that we can easily create a strict ordering $<_R$ from $\leq_R$ and vice versa. For this reason, we will not define properties for both strict and non-strict orderings separately, because one implicitly defines the other.

Definition 1.12. We call an element of an ordered class $(X, \leq)$ *minimal* if there is no smaller one, and we call it a *minimum* if it is smaller than all others. If a minimum exists, we denote it by $\min_{\leq}(X)$. The *supremum* of a subset $Y \subseteq X$ is the minimum of all its upper bounds. If it exists, we denote it by $\sup_{\leq}(Y)$

Observation 1.13. *Every minimum is minimal. Furthermore, if $\leq_R$ is a total order, then there is at most one minimal element, and if it exists, then it is also the minimum. There is always at most one minimum.*

Definition 1.14 (Well-ordering)**.** An ordered class $(A, \leq_R)$ is

(a) *well-founded* if every non-empty subset of A has a minimal element.

(b) *well-ordered* if every non-empty subset of A has a minimum (least element).

Notice that every well-ordered class is totally ordered since we can take any two elements, and one of them has to be the minimum and is therefore smaller.

Observation 1.15. *Well-order $\iff$ well-founded total order.*

Observation 1.16. *The well-ordering property is hereditary. That is, if X is well-ordered by $\leq_R$, then every $Y \subseteq X$ is also well-ordered by $\leq_R$.*

Observation 1.17. *Well-founded ordered sets contain no infinite strictly decreasing sequences, as such a sequence has no minimal element.*

Exercise 2. Proof that the reverse implication also holds, provided we accept the axiom of dependent choice (see Axiom 1.5).

Definition 1.18 (Lower part and subset)**.** Let $(A, <_R)$ be a (strictly) ordered class. A subclass $X \subseteq A$ is a *lower part* of A if

$$(\forall x \in X)(\forall a \in A)(a <_R x \Rightarrow a \in X).$$

Additionally, if X is a set, we call it a *lower subset* of A, and if $X \neq A$, then we call it a *proper lower part*, or *proper lower subset* of A.

Lemma 1.19. *Let $(W, <_R)$ be a (strictly) well-ordered set, and suppose that X is a proper lower subset of W. Then there exists a unique $x \in W$ such that X is equal to the set $\{y \in W \mid y <_R x\}$. We denote this set as $(\leftarrow, x)$.*

Proof. We define x as the minimum of $W \setminus X$. Then every $y <_R x$ belongs to X, so $(\leftarrow, x) \subseteq X$. We also want the opposite inclusion. For contradiction, suppose there is a $y \in X$ such that $y \notin (\leftarrow, x)$. If $y \not< x$, then necessarily $x \leq_R y$. But this means that $x \in X$ because X is a lower subset and $y \in X$. But this is a contradiction since $x \notin X$. □

Definition 1.20 (Initial segment)**.** If $(W, <_R)$ is a (strictly) well-ordered set, then we call its proper lower subsets *initial segments* instead. We denote the unique initial segment of W determined by $x \in W$ as

$$(\leftarrow, x) \coloneqq \{y \in W \mid y <_R x\}.$$

It contains all the elements of W from the minimum of W until x, but not x itself.

Observation 1.21. *Note that $x <_R y \iff (\leftarrow, x) \subset (\leftarrow, y)$.*

2 Ordinal Numbers

Informally, *ordinal numbers* are a way to generalize natural numbers. We will first do a quick recap of the basics of ordinal numbers and then prove a theorem that deeply links ordinals and well-ordered sets.

2.1 Ordinals as a Generalization of Naturals

Definition 2.1. A class X is called *transitive* if for all $x \in X$ we have $x \subseteq X$. Or equivalently, if for every x, y such that $y \in x \in X$ we have $y \in X$.

Theorem 2.2. *Every natural number and the set of all natural numbers ω is transitive and (strictly) well-ordered by the membership relation $\in$.*

From now on, we will denote the (strictly) well-ordered set $(\omega, \in)$ as $(\omega, <)$ instead and write $n < m$ instead of $n \in m$ when talking about natural numbers.

Definition 2.3 (Ordinal numbers)**.** A set α is an *ordinal number* if it is transitive and (strictly) well-ordered by the membership relation $\in$. If α is infinite, we say that it is a *transfinite ordinal*. We denote the *class of all ordinal numbers* by On.

Theorem 2.4. *Finite ordinals are exactly the natural numbers, and ω is the smallest transfinite ordinal.*

Theorem 2.5. *The class* On *itself is transitive and (strictly) well-ordered by $\in$. This implies that it is not a set; otherwise,* $\mathrm{On} \in \mathrm{On}$*. Furthermore, any proper class X that is transitive and well-ordered by $\in$ is identical to* On*.*

As for notation, we will usually denote ordinals using letters from the beginning of the Greek alphabet: $\alpha, \beta, \gamma, \delta \ldots$ An exception to this is the letter λ, which we will reserve for limit ordinals. Furthermore, we compare ordinals using the symbol '$<$'. That is, we write $\beta < \alpha$ instead of $\beta \in \alpha$.

Observation 2.6. *If $\beta < \alpha$, then $\beta \subset \alpha$ and β is an initial segment of α. Additionally, $\alpha = (\leftarrow, \alpha)$.*

Definition 2.7. If α is an ordinal, then we call all $\beta < \alpha$ the *predecessors* of α. The *successor* of α is the ordinal $\alpha^+ := \alpha \cup \{\alpha\}$. We say that α is the *direct predecessor* of α^+.

Remark. It is easy to show that α^+ is the smallest ordinal larger than α.

Definition 2.8. We say that an ordinal number α is an

(a) *isolated* ordinal if $\alpha = 0$ or α has a direct predecessor,

(b) a *limit* ordinal otherwise.

Isolated ordinals $\alpha > 0$ are also sometimes called *successor* ordinals.

Example. Every $n \in \omega$ is isolated, ω is limit, and ω^+ is isolated again.

2.2 Ordinals as Types of Well-Ordered Sets

The definition of ordinals presented above was formalized by John von Neumann in 1923. This elegant approach, however, came decades after Georg Cantor first introduced ordinals (around 1885) as *order types of well-ordered sets*. Cantor's intuition was that ordinals serve as labels for well-ordered sets: the smallest element is labeled 0, the next 1, and so on. The *order type* of the set is then the first label we did not have to use; it represents the "shape" of the ordering.

Consider, for example, a set ordered as

$$a_0 < a_1 < a_2 < \overbrace{\cdots}^{\infty} < b.$$

Here, there are countably infinitely many elements a_i, followed by one additional element b. If we label the elements from left to right, all the natural numbers are used for the a_i's, leaving no finite label for b. This is precisely why we need transfinite ordinals: we assign the label ω to b. Hence, the order type of this ordering is $\omega^+ = \omega + 1$.

It is important to realize that different orderings of the same sets can have different order types. This means that the ordinal numbers do not count the number of objects in the set; they only label them.

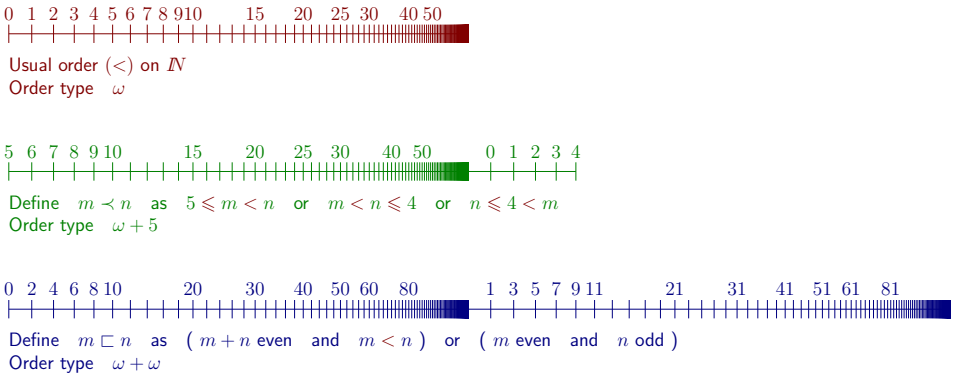

Figure 1: Different orderings of the same set can have different types; [5].

Lemma 2.9. *Every proper lower part of* (On, <) *is an ordinal number.*

Proof. Let X be a proper lower part of On. Then

(i) X is transitive. Suppose $\alpha \in \beta \in X$, that is $\alpha < \beta \in X$. Because X is a lower part, we have $\alpha \in X$.

(ii) X is well-ordered by $\in$ because On is well-ordered by $\in$ and $X \subseteq$ On.

We also need to argue that X is a set. If it were a proper class, then by Theorem 2.5 it would be the entire On, but $X \subsetneq$ On. □

Definition 2.10 (Isomorphism)**.** Let $(A, \leq_R)$ and $(B, \leq_S)$ be ordered classes. A bijection $F : A \to B$ is an *order-isomorphism* of $(A, \leq_R)$ and $(B, \leq_S)$ if

$$(\forall x,y \in A)\big(x \leq_R y \iff F(x) \leq_S F(y)\big).$$

Because we will not be dealing with other types of isomorphisms, we will usually simply say *isomorphism* instead of order-isomorphism.

Theorem 2.11 (About comparing well-orderings)**.** *If* $(W_1, \leq_1)$ *and* $(W_2, \leq_2)$ *are well-ordered sets, then exactly one of the following holds:*

(a) *either* W_1 *and* W_2 *are isomorphic, or*

(b) W_1 *is isomorphic to an initial segment of* W_2*, or*

(c) W_2 *is isomorphic to an initial segment of* W_1*.*

In each case, the isomorphism is unique.

Corollary 2.12. *No two distinct ordinal numbers can be isomorphic.*

Proof. Suppose $\alpha < \beta$, that is $\alpha \in \beta$ and $\alpha \subset \beta$. Clearly α is an initial segment of β. This means that we are in case (b) of the previous theorem. □

Theorem 2.13 (About the type of well-ordering)**.** *Every well-ordered set is isomorphic to a unique ordinal number, which is called the* type *of the ordering.*

Proof (following [16, Thm. 3.1, Chap. 6]*).* Let $(W, <_R)$ be a well-ordered set. We want to show that there is a unique ordinal $(\alpha, <)$ isomorphic to it. Define X as the set of all $x \in W$ for which $(\leftarrow, x)$ is isomorphic to an ordinal. As no two distinct ordinals are isomorphic, this ordinal is uniquely determined, and we denote it α_x; we denote the isomorphism as $i_x : (\leftarrow, x) \to \alpha_x$.

Suppose that there exists a set S such that $S = \{\alpha_x \,|\, x \in X\} \subseteq \mathrm{On}$. Because we assume that S is a set, then $S \subsetneq \mathrm{On}$. We claim that S is a proper lower part of $(\mathrm{On}, <)$, and thus, by Lemma 2.9, it is an ordinal; let's call it α. Indeed, suppose $\beta < \alpha_x \in S$, we want $\beta \in S$. Note that β is an initial segment of α_x. This implies that $i_x^{-1}[\beta]$ is an initial segment of W. Because W is well-ordered, $i_x^{-1}[\beta]$ is equal to $(\leftarrow, b)$ for some $b \in W$ (using Lemma 1.19). So $\beta = \alpha_b \in S$ by the definition of S. More precisely, $i_x \restriction (\leftarrow, b)$ is an isomorphism between $(\leftarrow, b)$ and β. We will argue why we can make the assumption that S is a set later.

A similar argument shows that X is a lower subset of W. To show this, suppose $x \in X$ and take $y \in a$ such that $y <_R x$. We want $y \in X$. We have $y <_R x$; therefore, $(\leftarrow, y)$ is an initial segment of $(\leftarrow, x)$. Because isomorphisms conserve all ordering properties, $i_x \restriction (\leftarrow, y)$ is an isomorphism between $(\leftarrow, y)$ and an initial segment of α_x. By Lemma 2.9, this is an ordinal; by our previous notation, α_y. Therefore $y \in X$.

We conclude that either $X = W$ or $X = (\leftarrow, c) \subset W$ for some $c \in W$ (using Lemma 1.19). We now define a function $f : X \to S = \alpha$ by $f : x \mapsto \alpha_x$. From the definition of S and the fact that $x < y$ implies $(\leftarrow, x) \subset (\leftarrow, y)$ and therefore $\alpha_x < \alpha_y$, it is obvious that f is an isomorphism of $(X, <_R)$ and $(\alpha, <)$. If

- $X = (\leftarrow, c)$, then by the definition of the set X, $c \in X$ because $(\leftarrow, c)$ is isomorphic to an ordinal $\alpha_c = \alpha$. But this is a contradiction because $c \notin (\leftarrow, c) = X$.

- Therefore, $X = W$ and α is the sought-after ordinal isomorphic to $(W, <_R)$.

The uniqueness of α follows from the simple observation that if W were isomorphic to two distinct α_1 and α_2, then by the transitivity of isomorphisms, the ordinals α_1 and α_2 would be isomorphic, which is impossible by Corollary 2.12.

This would complete the proof if we were justified to make the assumption that the class S is a set and therefore an ordinal. In fact, we have to use the axiom of replacement to guarantee it. If we assume this axiom, then S is a set because it is the image of the set X under the map f. □

Exercise 3. Is there a well-ordered proper class not isomorphic to (On, <)?

Hint. Try to modify On to contradict the property described in Lemma 2.9. If you run out of ideas, consult Section 3.2.

2.3 Transfinite Induction and Recursion

In mathematics, we often use induction on the natural numbers to prove statements, and we can use recursion, such as $f(0) = 1$ and $f(n) = n \cdot f(n-1)$, to define functions. We will now show how to generalize this to all ordinals.

Theorem 2.14 (Transfinite Induction Principle)**.** *Let $A \subseteq \mathrm{On}$ be a class such that for all ordinals $\alpha \in \mathrm{On}$, we have $\alpha \subseteq A \Rightarrow \alpha \in A$, or in other words*

$$(\forall \beta < \alpha)(\beta \in A) \implies (\alpha \in A). \tag{2.1}$$

Then $A = \mathrm{On}$.

Equivalently, assume that $\varphi(x)$ is a property, and for all ordinals α:

If $\varphi(\beta)$ holds for all $\beta < \alpha$, then $\varphi(\alpha)$.

Then $\varphi(\alpha)$ holds for all ordinals α.

Proof. Suppose that $\gamma \in \mathrm{On} \setminus A$ and let $S = \{\alpha \leq \gamma \mid \alpha \notin A\}$. Because ordinals are well-ordered, the set S has a minimum element α. Since every $\beta < \alpha$ is in A, it follows by (2.1) that $\alpha \in A$, which is a contradiction.

The equivalence can be easily seen by taking the class $A = \{x \mid \varphi(x)\}$ or the property $\varphi(x) = x \in A$. □

We can also formulate the principle separately for isolated and limit ordinals, which allows us to use the transfinite induction principle in a form closer to the usual formulation of the induction principle for the naturals.

Theorem 2.15 (Transfinite Induction Principle II)**.** *Let $A \subseteq \mathrm{On}$ be a class satisfying*

(i) $0 \in A$,

(ii) $\alpha \in A \Rightarrow \alpha^+ \in A$, *... this is just induction on ω*

(iii) *if α is a limit ordinal and $(\forall \beta < \alpha)(\beta \in A)$, then $\alpha \in A$.*

Then $A = \mathrm{On}$. Note that we can again easily reformulate this in terms of a property $\varphi(x)$.

Proof. We need to show that these three assumptions imply (2.1). So let α be an ordinal such that $\beta \in A$ for all $\beta < \alpha$. If $\alpha = 0$, then $\alpha \in A$ by (i). If $\alpha \neq 0$ is isolated, that is if there is a $\beta < \alpha$ such that $\alpha = \beta^+$, we know that $\beta \in A$, so $\alpha \in A$ by (ii). If α is a limit ordinal, we have $\alpha \in A$ by (iii). □

We can use transfinite induction to prove properties of certain infinite structures. On the other hand, transfinite recursion—the technique described in the following theorem—allows us to construct various infinitely complex structures and define functions in a recurrent fashion.

Theorem 2.16 (About construction by transfinite recursion)**.** *If $G : V \to V$ is a class map, then there is a unique class map $F : \mathrm{On} \to V$ satisfying*

$$F(\alpha) = G(F \restriction \alpha). \tag{2.2}$$

So we define the image of the next ordinal using its predecessors and their images.

Remark. This should seem a bit suspicious because it looks like we are saying that for every class G, there exists a class F for which something holds. But we cannot quantify classes. Well, we can replace the quantification of G with a theorem *schema*, one for each G. And we aren't really quantifying F because the following proof explicitly constructs it.

Remark. The theorem can be equivalently formulated using different recurrences; for example, as

- $F(\alpha) = G\big(F[\alpha]\big) = G\big(\{F(\beta) \mid \beta < \alpha\}\big)$,
- $G : \mathrm{On} \times V \to V$ and $F(\alpha) = G(\alpha, F \restriction \alpha)$,
- $F(\alpha)$ is $G_1(F(\beta))$ if $\alpha = \beta^+$ is isolated, and $G_2(F[\alpha])$ if α is limit.

Additionally, these transfinite recursion statements are equivalent to the axiom of replacement, as shown in [15].

Proof. We define A as the class of "set approximations" of F. That is set mappings f, the domain of which is some ordinal number β, and that for all $\alpha < \beta$, we have $f(\alpha) = G(f \restriction \alpha)$. Now we define F as $F \coloneqq \bigcup A$. Clearly $F \subseteq \mathrm{On} \times V$. We will show that $F : \mathrm{On} \to V$ is the unique mapping satisfying (2.2).

First, we show that the approximations of F agree. Let f,$f' \in A$ and $\alpha \in \mathrm{Dom}(f) \cap \mathrm{Dom}(f')$. We claim that $f(\alpha) = f'(\alpha)$. Note that $\mathrm{Dom}(f) \cap \mathrm{Dom}(f')$ is an ordinal δ. For contradiction, suppose that $\alpha \in \delta$ is the smallest ordinal for which $f(\alpha) \neq f'(\alpha)$. Then $f \restriction \alpha = f' \restriction \alpha$ so $f(\alpha) = G(f \restriction \alpha) = G(f' \restriction \alpha) = f'(\alpha)$, a contradiction.

Second, we verify that F satisfies (2.2); that is, for all $\alpha \in \mathrm{Dom}(F)$, we have $F(\alpha) = G(F \restriction \alpha)$. So let $\alpha \in \mathrm{Dom}(F)$. It is there due to some $f \in A$ satisfying $\alpha \in \mathrm{Dom}(f)$ and $f(\alpha) = G(f \restriction \alpha)$. Also, $F(\alpha) = f(\alpha)$ and $F \restriction \alpha = f \restriction \alpha$. Therefore, by combining these equalities $F(\alpha) = G(F \restriction \alpha)$.

Next, we show that $\mathrm{Dom}(F) = \mathrm{On}$. First, we prove that $\mathrm{Dom}(F)$ is a lower part of On. Suppose $\alpha \in \mathrm{Dom}(F)$; then it is there thanks to some $f \in A$ with domain $\delta > \alpha$. If $\beta < \alpha$, then also $\beta \in \delta$, and thus $\beta \in \mathrm{Dom}(F)$.

According to Lemma 2.9, either $\mathrm{Dom}(F) = \mathrm{On}$, which we want, or $\mathrm{Dom}(F) = \gamma \in \mathrm{On}$. Suppose, for contradiction, that $\mathrm{Dom}(F) = \gamma$. Then F is a set because

$\mathrm{Dom}(F)$ is a set, $\mathrm{Rng}(F)$ is a set using the axiom of replacement, and $F \subseteq \mathrm{Dom}(f) \times \mathrm{Rng}(f)$. This implies that $F \in A$ because its domain is an ordinal, and we have verified that it satisfies the recursive definition property.

Now that $F \in A$, we define a slightly "longer" function $F_1 \coloneqq F \cup \{(\gamma, G(F))\}$; note that $F = F_1 \restriction \gamma$. Notice that $F_1 \in A$ because $\mathrm{Dom}(F_1) = \gamma^+$ is an ordinal, and we defined it to satisfy the recursive definition property. Because $F = \bigcup A$, this implies $F_1 \subseteq F$, but then $\gamma \in \mathrm{Dom}(F_1) \subseteq \mathrm{Dom}(F) = \gamma$, which is a contradiction. We conclude that $\mathrm{Dom}(F) = \mathrm{On}$.

Finally, we prove the uniqueness of F. For contradiction, suppose that there is another mapping $F' \neq F$ satisfying this theorem. Because $(\mathrm{On}, <)$ is well-ordered, we can take the smallest ordinal α where $F(\alpha) \neq F'(\alpha)$. Therefore $F \restriction \alpha = F' \restriction \alpha$ and so $F(\alpha) = G(F \restriction \alpha) = G(F' \restriction \alpha) = F'(\alpha)$, which is a contradiction. □

Exercise 4. Prove by induction on ω that every infinite well-ordered set A, such that each initial segment $(\leftarrow, a)$ is finite, is isomorphic to $(\omega, <)$.

Hint. Since each $(\leftarrow, a)$ is finite, there is a unique $n_a \in \omega$ with the same cardinality. The isomorphism we are looking for is $f : A \to \omega$ defined by $f : a \mapsto n_a$.

Exercise 5. Prove by transfinite recursion that every well-ordered proper class W, such that each proper lower part $(\leftarrow, a)$ is a set, is isomorphic to $(\mathrm{On}, <)$.

Hint. Use transfinite recursion to define an isomorphism $F : \mathrm{On} \to W$ using $G(x) = \min(W \setminus x)$ as $F(0) = \min(W)$ and $F(\alpha) = G(F[\alpha])$.

We will use transfinite recursion to prove the equivalence of the well-ordering theorem and Zorn's lemma to the axiom of choice. But transfinite recursion can also be used to prove some wildly sounding geometrical claims, such as

- $\mathbb{R}^3$ is a union of pair-wise disjoint unit circles, or that
- there is a set in $\mathbb{R}^2$ that intersects every line in exactly two points.

2.4 The Well-Ordering Principle

The *well-ordering principle*—the statement that every set can be well-ordered—was a foundational belief of Georg Cantor, but he was unable to provide a proof for it. This challenge was famously solved by Ernst Zermelo in 1904. Zermelo was the first person to explicitly state the axiom of choice, which he identified as the principle Cantor (and many others) had been implicitly using in many proofs. He then demonstrated that AC and the well-ordering principle are equivalent, which is why the principle is now often called the "Well-Ordering Theorem" or "Zermelo's Theorem." Veritasium has a great video [25] on this topic.

Principle 2.17 (Well-Ordering Principle)**.** Every set can be well-ordered.

Theorem 2.18. *The well-ordering principle is equivalent to the axiom of choice.*

Proof. $\mathsf{WO} \Rightarrow \mathsf{AC}$. Let $A \neq \varnothing$ be a set, without loss of generality $\varnothing \notin A$. We want to construct a selector $f : A \to \bigcup A$ such that for all $a \in A$ we have $f(a) \in a$. The well-ordering principle guarantees a well-ordering $\leq$ on $\bigcup A$, and

because every a is a nonempty subset of $\bigcup A$, it has a least element with respect to $\leq$. We chose this minimum as $f(a)$.

$\mathsf{AC} \Rightarrow \mathsf{WO}$. Let $A \neq \varnothing$ be a set. We will use transfinite recursion to label the elements of A by ordinal numbers and then use the well-order of the ordinals to define a well-order on A. Let $g : \mathcal{P}(A) \to A$ be a selector on $\mathcal{P}(A)$, assigning to each nonempty $B \subseteq A$ an element $b \in B$. We will want to use transfinite recursion based on g, so we should extend it to be a class map $G : V \to V$, for example, by defining it to be equal to $\varnothing$ when g is not defined.

We can now use transfinite recursion to define the function $F : \mathrm{On} \to A \cup \{\varnothing\}$ as $F(0) = G(A)$ and $F(\alpha) = G(A \setminus F[\alpha])$. This function assigns to each ordinal a unique element from A until they "run out" (when $F[\alpha] = A$), and then it assigns $\varnothing$ to all larger ordinals.

Define W as the class of all ordinals α for which $F[\alpha] \subsetneq A$. Denote the restriction of F to W as $F_W : W \to A$. Plan: first, we show that W itself is an ordinal. From this, it will follow that F_W is a bijection between W and A, allowing us to denote the unique ordinal mapped to $a \in A$ as α_a. Once this is established, we define a well-ordering R of A as

$$a <_R b \iff \alpha_a < \alpha_b.$$

This is a well-ordering since $(A, <_R)$ is order-isomorphic to $(W, <)$, which is well-ordered (as W is an ordinal).

Firstly, we claim that W is a set. Indeed, because F_W is injective, it has an inverse F_W^{-1} that maps the set $\mathrm{Rng}(F_W) \subseteq A$ onto W, which is therefore, using the axiom of replacement, a set. Now we claim that W is a lower subset of On, and so it is an ordinal (by Lemma 2.9). Suppose $\alpha \in W$, that is $F[\alpha] \subsetneq A$, and let $\beta < \alpha$. Then $\beta \subseteq \alpha$ and $F[\beta] \subseteq F[\alpha]$, so $\beta \in W$.

To complete the proof, we must show that $F_W : W \to A$ is a bijection. It is clearly injective. To show that it is surjective, suppose for contradiction that there exists some $b \in A \setminus F_W[W]$. Because W is an ordinal number γ, it satisfies the definition of W (thanks to b) and thus $W = \gamma \in W$, which is a contradiction. □

2.5 Zorn's Lemma

Zorn's lemma is perhaps the most useful application of the axiom of choice outside of set theory. It is also known as the maximality principle, a name that dates back to the German mathematician Felix Hausdorff, who proved an earlier and equivalent version of the theorem in 1914 (see [28] for details). The formulation known today as Zorn's lemma was introduced in 1935 by the German mathematician Max Zorn. However, it had already been independently proven in 1922 by the Polish mathematician Kazimierz Kuratowski, whom you might know for Kuratowski's theorem—a forbidden-graph characterization of planar graphs.

Definition 2.19 (Chain)**.** Let $(a, \leq_R)$ be an ordered set. We call the subset $b \subseteq a$ a *chain* if b is totally ordered by $\leq_R$.

Principle 2.20 (Zorn's Lemma)**.** Every (partially) ordered set containing upper bounds for every chain necessarily contains at least one maximal element.

There is also a parameterized version of this statement.

Principle 2.21 (Parametrized Zorn's Lemma)**.** Let A be a (partially) ordered set containing upper bounds for every chain. Then for every $a \in A$, there is a maximal element $b \in A$ such that $a \leq b$.

We can obtain the parameterized version from the unparameterized one by restricting ourselves to the elements above or equal to a. The other direction is obvious.

Remark. Zorn's lemma can be made slightly stronger by assuming that only well-ordered chains have upper bounds. The proof remains virtually unchanged.

Theorem 2.22. *The axiom of choice implies Zorn's lemma.*

Proof. Let $(A, <_R)$ be an ordered set containing upper bounds for each chain and for contradiction suppose that there is no maximal element. Note that this implies that every chain, in fact, has a *strict* upper bound. If a chain C had no strict upper bound, then the non-strict upper bound $b \in C$ would be a maximal element. We denote the set of strict upper bounds of C as $C^>$.

We take $f : \mathcal{P}(A) \to A$, a selector on $\mathcal{P}(A)$, and define a function g from the set of all chains in A as $g(C) \coloneqq f(C^>)$. So g maps a chain to one of its strict upper bounds. Now pick an arbitrary $a \in A$ and define the mapping $H : \mathrm{On} \to A$ by transfinite recursion as $H(0) = a$ and $H(\alpha^+) = g(\{H(\alpha)\})$ for successor ordinals, and as $H(\delta) = g(H[\delta])$ for limit ordinals. We start with a and get larger and larger elements of A using successor ordinals, each time taking a strict upper bound of a single element chain. If an ordinal δ is limit, we notice that $H[\delta]$ is a chain (all the smaller elements that we picked previously are strict upper bounds of each other and are therefore comparable), and $H(\delta)$ is a strict upper bound of this chain.

Note that if we want to be rigorous about the construction by transfinite recursion, we should define g on the entire V. But we can do this in any way, for example, by defining $G(x)$ as $\varnothing$ if x is not a chain of A, and $g(x)$ otherwise.

Finally, observe that $H : \mathrm{On} \to A$ is an increasing function (each value is a strictly larger upper bound than the previous one) and that it is injective. Thus, we obtain an injection from the proper class On into the set A, which is impossible. Indeed, taking the inverse mapping and applying the axiom of replacement would imply that On itself is a set, which is a contradiction. □

Theorem 2.23. *Zorn's lemma implies the well-ordering principle.*

Proof. Let X be any set. We will find a well-ordering of it by considering all of its possible well-ordered subsets, picking the maximal one using Zorn's lemma, and showing that it orders the entire X. Consider the set: [4]

$$\mathcal{W} \coloneqq \{(A, <_R) \mid <_R \text{ is a well-order on } A \subseteq X\},$$

and define a partial order $\prec_{\mathcal{W}}$ on it by $(A, <_R) \prec_{\mathcal{W}} (B, <_S)$ if B end-extends A. That is, if $A \subset B$, and $<_R$ is the restriction of $<_S$ to A, and A is an initial segment of B. We will apply Zorn's lemma to $\mathcal{W}$.

First, we need to show that chains have upper bounds. Let $\mathcal{C} \subseteq \mathcal{W}$ be a chain. Define the set

$$M \coloneqq \bigcup \{A \mid (A, <_R) \in \mathcal{C}\} \subseteq X,$$

[4] Why is this a set?

and for $x, y \in M$ put $x <_M y$ if there exists some $(A, <_R) \in \mathcal{C}$ such that $x, y \in A$ and $x <_R y$. Because $\mathcal{C}$ is a chain, this is well-defined: if x and y belong to two distinct orderings in $\mathcal{C}$, then one extends the other and hence they agree.

We claim that $(M, <_M)$ is well-ordered. Let $S \subseteq M$ be nonempty and pick some $s \in S$. Then $s \in A_s$ for some $(A_s, <_R) \in \mathcal{C}$. Note that $A_s \cap S$ is nonempty, and because A_s is well-ordered, there exists a minimum $m = \min_{<_R}(A_s \cap S)$. Notice that also $m = \min_{<_M}(S)$. Indeed, if there were a $t \in S \setminus A_s$ such that $t <_M m$, then it would be in S due to some $A_t \in \mathcal{C}$ containing t. Since both A_s and A_t are in the chain, either

(a) $A_t \subseteq A_s$, which is impossible since then $t \in A_s$, contradicting the minimality of m in $A_s \cap S$, or

(b) $A_s \subset A_t$, meaning that A_s is an initial segment of A_t, and therefore $m \in A_s$ is smaller than $t \in A_t \setminus A_s$, which contradicts the assumption that $t <_M m$.

Therefore $(M, <_M)$ is well-ordered and thus an upper bound of $\mathcal{C}$ in $\mathcal{W}$.

Because all chains are bounded, by Zorn's lemma, $\mathcal{W}$ has a maximal element $(W, <_W)$. We claim that $W = X$ and so it is the sought-after well-ordering of X. For contradiction, suppose there exists some $x \in X \setminus W$ and extend the ordering $<_W$ to $W' := W \cup \{x\}$ by making each $y \in W$ smaller than x. Notice that this slightly "longer" order is a well-ordering of W' and therefore is in $\mathcal{W}$. Moreover, it end-extends $(W, <_W)$ which hence is not maximal in $(\mathcal{W}, \prec_{\mathcal{W}})$. We have arrived at a contradiction and can conclude that $W = X$. □

Exercise 6. Would the proof still have worked if instead of end-extensions, we had simply used general extensions? Meaning that the smaller ordering doesn't need to be an initial segment of the larger one.

Hint. By defining the end-extension ordering, we have ensured that chains have a similar structure to chains of ordinals (larger ordinals end-extend the smaller ones). Thus, when proving that M is well-ordered, we could have used a similar strategy as when proving that the ordinals are well-ordered.

To demonstrate an application of Zorn's lemma, consider the following question. Does every connected graph have a spanning tree? Finding one in a finite graph is easy: simply remove the edges of cycles until there are no cycles left. But this process may not terminate for infinite graphs.

Proposition 2.24. *Every connected graph has a spanning tree.*

Sketch of proof. The set of all sub-graphs that are trees is partially ordered by inclusion, and the union of a chain is its upper bound. Zorn's lemma states that a maximal tree must exist, which is a spanning tree since the graph is connected. □

Remark. In general, suppose that we have a structure represented by a set X (a graph) with substructures $A \subseteq X$ (subgraphs that are trees), and we want to show that there is a maximal substructure. Then we simply need to check that the union of a chain of substructures is itself a substructure.

2.6 The Trichotomy Principle

Definition 2.25. For sets x and y we define the relations

(a) $x \approx y$, if there exists a bijection $x \to y$,

(b) $x \preceq y$, if there exists a injection $x \to y$,

(c) $x \prec y$, if $x \preceq y$ and $x \not\approx y$.

Definition 2.26. We say that a set x is

(a) *countable* if $x \preceq \omega$,

(b) *countably infinite* if $x \approx \omega$,

(c) *uncountable* if $\omega \prec x$.

Theorem 2.27 (Cantor, Berstein)**.** $x \approx y \iff (x \preceq y \wedge y \preceq x)$.

Principle 2.28 (Trichotomy principle)**.** The relation $\preceq$ is is trichotomous on V. That is, for any sets x and y either $x \preceq y$, or $y \preceq x$.

Theorem 2.29. *Zorn's lemma implies the trichotomy principle.*

Proof. Let x, y be arbitrary sets; we want to find an injection $x \to y$ or $y \to x$. Consider the set[5]

$$\mathcal{F} = \{f \mid f \text{ is an injection, } \mathrm{Dom}(f) \subseteq x \text{ and } \mathrm{Rng}(f) \subseteq y\}.$$

Notice that the ordered set $(\mathcal{F}, \subseteq)$ satisfies the conditions of Zorn's lemma since the union of a chain of injections is again an injection. Let g be a maximal element of $\mathcal{F}$. If both $x \setminus \mathrm{Dom}(g)$ and $y \setminus \mathrm{Rng}(g)$ were non-empty, then it would be possible to extend g by an extra pair, contradicting its maximality. Hence either $\mathrm{Dom}(g) = x$ and then $x \preceq y$, or $\mathrm{Rng}(g) = y$ and then $y \preceq x$. Here, we used the fact that the inverse of an injection is also an injection. □

Later, (Theorem 3.47), we will show that the trichotomy principle implies the well-ordering principle and thus also the axiom of choice.

Theorem 2.30. *We conclude that the following statements are equivalent in* ZF*:*

- *the axiom of choice,*
- *the well-ordering principle,*
- *Zorn's lemma,*
- *the trichotomy principle.*

[5] Why is this a set?

3 Operations on Ordinals

3.1 Ordinal Functions

Definition 3.1 (Ordinal function)**.** We say that a mapping F is an *ordinal function* if its domain is a lower part of On, that is $\mathrm{Dom}(F) \in \mathrm{On}$ or $\mathrm{Dom}(F) = \mathrm{On}$, and $\mathrm{Rng}(F) \subseteq \mathrm{On}$. We say that F is

(a) *increasing* if for all $\beta \in \mathrm{Dom}(F)$ and $\alpha < \beta$ we have $F(\alpha) < F(\beta)$,

(b) *nondecreasing* if for all $\beta \in \mathrm{Dom}(F)$ and $\alpha < \beta$ we have $F(\alpha) \leq F(\beta)$.

Remark. We dont define a decreasing ordinal function because they aren't very interesting — F can be decreasing only when $\mathrm{Dom}(F)$ is finite, since the well-ordering On does not allow infinite decreasing sequences (see Observation 1.17).

Lemma 3.2. *Increasing ordinal functions grow at least as fast as the identity function. That is $F(\alpha) \geq \alpha$ for every $\alpha \in \mathrm{Dom}(F)$ for increasing F.*

Proof. For contradiction, suppose that α is the least ordinal such that $F(\alpha) < \alpha$. This means that for every $\beta < \alpha$, we have $F(\beta) \geq \beta$ (note that $\beta \in \mathrm{Dom}(F)$). Suppose $\beta = F(\alpha)$; then $F(\beta) \geq \beta = F(\alpha)$, which is a contradiction since F is increasing. □

Lemma 3.3. *If α and β are the ordinal types of the well-ordered sets A and $B \subseteq A$, then $\alpha \leq \beta$.*

Note that not necessarily $\alpha < \beta$ when $B \subsetneq A$; consider ω and $\omega \setminus \{\varnothing\}$.

Proof. Let $i_a : A \to \alpha$ and $i_b : B \to \beta$ be the isomorphisms of A and B with their types. Suppose $\beta > \alpha$ and define $f : \beta \to \alpha$ as $f = i_b^{-1} \circ i_a$. Notice that if $\gamma < \delta$, then $f(\gamma) < f(\delta)$ because both i_b^{-1} and i_a preserve order (they are order-isomorphisms), and thus f is increasing. Because $\alpha \in \mathrm{Dom}(f)$, we have that $f(\alpha) \in \mathrm{Rng}(f) = \alpha$. But this contradicts the previous lemma. □

Recall what you know about metric spaces, namely about closed sets and continuous functions. A subset X of a metric space M is closed if for each convergent sequence $(a_n) \subset X$ we have that $\lim a_n \in X$.

You might also recall that a function f is continuous $\iff$ the preimage $f^{-1}[Y]$ of every closed set Y is closed $\iff$ for every sequence (a_n) we have that $f(\lim a_n) = \lim f(a_n)$. The first equivalence statement is utilized in topology to define continuous functions, and we could use it here as well. However, the second equivalence seems more natural, since $\lambda = \sup\{\alpha \mid \alpha < \lambda\}$ for any limit ordinal λ. Hence limit ordinals essentially represent limits of sequences.

Lemma 3.4. *If $A \subseteq \mathrm{On}$ is a set, then $\bigcup A \in \mathrm{On}$, and in fact, $\bigcup A = \sup(A)$. We say that $\sup(A)$ is the* limit *of the sequence of ordinals A.*

Definition 3.5. A subclass $C \subseteq \mathrm{On}$ is *closed* if, for every subset $Y \subseteq C$, we have $\sup(Y) \in C$. For an ordinal α, we say that a subset $C \subseteq \alpha$ is *closed in α* if, for every $Y \subseteq C$ satisfying $\sup(Y) < \alpha$, we have $\sup(Y) \in C$.

Observation 3.6. *If C is closed, then it has a maximum $\max(C) = \sup(C)$.*

Example. The ordinal ω is not closed because $\sup(\omega) = \omega \notin \omega$, but ω^+ is closed. In general, isolated ordinals are closed and limit ordinals are not.

Definition 3.7. An ordinal function F is *continuous* if for every limit ordinal $\lambda \in \mathrm{Dom}(F)$ it holds that

$$F(\lambda) = \sup\{F(\alpha) \mid \alpha < \lambda\}.$$

We say that a function is *normal* if it is increasing and continuous.

Example. The simplest normal function is identity. But consider the (very innocent looking) function $F(\alpha) = \alpha^+$. It is increasing, but not continuous. It fails on limit ordinals, for example $F(\omega) = \omega^+$, but

$$\sup\{F(\alpha) \mid \alpha < \omega\} = \sup\{\alpha^+ \mid \alpha < \omega\} = \sup(\omega \setminus \{\varnothing\}) = \omega.$$

Observation 3.8. *If F is a normal ordinal function and λ is a limit ordinal, then $F(\lambda)$ is a limit ordinal as well.*

Proof. Suppose that $F(\lambda) = \sup\{F(\alpha) \mid \alpha < \lambda\}$ were isolated. Then $F(\lambda) = \gamma^+$ for some $\gamma < \lambda$, which has been taken into account in the supremum. Therefore $F(\lambda) = F(\gamma)$, since F is increasing and γ is the largest ordinal smaller than λ. But this is a contradiction because $F(\gamma) < F(\lambda)$. □

Exercise 7. Show that the *topological definition of continuity* would make sense. Prove that an increasing function F is continuous $\iff$ the preimage $F^{-1}[C]$ of every closed set $C \subset \mathrm{On}$ is closed in $\mathrm{Dom}(F)$.

Observation 3.9. *The composition $F \circ G$ of normal functions F and G is normal. This can be seen easily from the topological definition of a continuous function.*

Lemma 3.10. *If F is a normal function, then for every ordinal β, such that $F(0) \leq \beta < \sup \mathrm{Rng}(F)$, the maximum $\max\{\alpha \mid F(\alpha) \leq \beta\}$ exists.*

Intuition. For a natural number β and $f(n) = n^2$, we might consider the largest natural number α such that $F(\alpha) \leq \beta$. This α exists, it is in fact equal to $\lfloor \sqrt{\beta} \rfloor$.

Proof. Consider the closed set $[0, \beta] \coloneqq \{\alpha \mid \alpha \leq \beta\}$. It is indeed closed because $[0, \beta] = \beta^+$, which is an isolated ordinal. We will use the topological definition of continuity (Exercise 7) and note that the preimage C of the closed set $[0, \beta]$ is closed in $\mathrm{Dom}(F)$. We would like to say that C is closed (in general). But consider $F : \omega \to \mathrm{On}$; then $\mathrm{Dom}(F)$ is not closed in On.

The bound on β will save us. Notice that there is some $\gamma \in \mathrm{Rng}(F)$ such that $\beta < \gamma \leq \sup \mathrm{Rng}(F)$. Because F is increasing, the elements of C are bounded by $F^{-1}(\gamma)$, and therefore $\sup(C) \in \mathrm{Dom}(F)$. Since C is closed in $\mathrm{Dom}(F)$, it follows that $\sup(C) \in C$ and $\sup(C) = \max(C)$, which we will denote as α. Because F is increasing, α is the largest ordinal satisfying $F(\alpha) \leq \beta$. □

Definition 3.11. An ordinal ξ is a *fixed point* of an ordinal function F if $F(\xi) = \xi$. The *class of all fixed points* of F is denoted by $K(F)$.

Example. The fixed points of the identity function are all ordinals, while the function we saw earlier, $f(\alpha) = \alpha^+$, has no fixed points. We will show that the reason is that it is not continuous.

Theorem 3.12 (About fixed points)**.** *Let $F : \mathrm{On} \to \mathrm{On}$ be a normal function.*

(i) *For every $\alpha \in \mathrm{On}$, there exists $\beta \geq \alpha$, which is a fixed point of F.*

(ii) *Concretely, consider the sequence $(\alpha_n \,|\, n \in \omega)$ defined as $\alpha_0 \coloneqq \alpha$ and $\alpha_{n+1} = F(\alpha_n)$. The supremum of this sequence is the smallest of all fixed points $\xi \geq \alpha$ of F.*

(iii) *The class $K(F)$ is closed and is a proper class.*

Proof. We begin by proving (i) and (ii). Notice that $\alpha_{n+1} \geq \alpha_n$ since F grows at least as fast as the identity function. First, we show that the supremum $\beta = \sup\{\alpha_n \,|\, n \in \omega\}$ is a fixed point:

- Consider the case when $\alpha_0 < \alpha_1 < \cdots < \alpha_i = \alpha_{i+i}$ for some i; then also $\alpha_{i+2} = F(\alpha_{i+1}) = F(\alpha_i) = \alpha_i$, and thus $\beta = \alpha_i$ is a fixed point.
- Suppose the sequence never stabilizes. Since F is continuous, we have

$$\begin{aligned} F(\beta) &= \sup\{F(\gamma) \,|\, \gamma < \beta\} = \sup\{F(\alpha_n) \,|\, n \in \omega\} \\ &= \sup\{\alpha_{n+1} \,|\, n \in \omega\} = \sup\{\alpha_n \,|\, n \in \omega\} = \beta. \end{aligned}$$

Second, we show that β is the smallest fixed point larger than α. If there were a fixed point $\xi \geq \alpha$ such that $\xi < \beta$, then there would exist an index n at which $\alpha_n \leq \xi < \alpha_{n+1}$. This is because the sequence is strictly increasing, and β is its supremum. We have $\xi < \alpha_{n+1} = F(\alpha_n) \leq F(\xi)$, so ξ isn't a fixed point.

Finally, we prove (iii). We claim that $K \coloneqq K(F)$ is closed. Let $C \subseteq K$ be a set; we need to show that the supremum $\beta = \sup(C)$ is a fixed point. Since F is continuous, we have

$$F(\beta) = \sup\{F(\gamma) \,|\, \gamma < \beta\} = \sup\{F(\xi) \,|\, \xi \in C\} = \sup\{\xi \,|\, \xi \in C\} = \beta.$$

To complete the proof, we show that K is a proper class. If it were a set, then by Lemma 3.4, $\sup(K)$ is an ordinal γ. We let the ordinal γ^+ take the role of α in (i) and find a new fixed point of F, larger than all those in K. □

Corollary 3.13. *If $F : \mathrm{On} \to \mathrm{On}$ is a normal function, then there exists a unique order-isomorphism $J : \mathrm{On} \to K(F)$. Moreover, J is a normal function.*

Remark. The function J is also sometimes referred to as the *derivative* of the normal function F and is denoted by F'.

Proof. The proper class $K(F) \subseteq \mathrm{On}$ is well-ordered, and it inherits from On the property that every proper lower part $(\leftarrow, a) \subset K(F)$ is a set (see Lemma 2.9). Exercise 5 claims that there is a unique isomorphism $J : \mathrm{On} \to K(F)$. The ordinal function J is clearly increasing (since it preserves order). It is continuous because $K(F)$ is closed. Let $\lambda \in \mathrm{On}$, we claim that $J(\lambda) = \sup\{J(\alpha) \,|\, \alpha < \lambda\}$. The set $\{J(\alpha) \,|\, \alpha < \lambda\}$ is a subset of $K(F)$, and thus its supremum lies in $K(F)$. Because λ is the smallest ordinal larger than every $\alpha < \lambda$, and J is an order-isomorphism, J has to map λ to the smallest $\kappa \in K(F)$ larger than all $\{J(\alpha) \,|\, \alpha < \lambda\}$. This is the above-mentioned supremum. □

Theorem 3.14 (About simultaneous fixed points)**.** *Let* $\langle F_i \mid i \in I \rangle$ *be a family of normal functions* $F_i : \mathrm{On} \to \mathrm{On}$ *indexed by a set* I*. We say that* ξ *is a* simultaneous fixed point *of* $\langle F_i \mid i \in I \rangle$ *if* $F_i(\xi) = \xi$ *for all* $i \in I$*.*

(i) *For every* $\alpha \in \mathrm{On}$*, there exists* $\beta \geq \alpha$*, which is a simultaneous fixed point of* $\langle F_i \mid i \in I \rangle$*.*

(ii) *The first simultaneous fixed point* $\geq \alpha$ *is the limit of the sequence* $\alpha_0 = \alpha$ *and* $\alpha_{n+1} = \sup\{F_i(\alpha_n) \mid i \in I\}$*.*

(iii) *The class* K *of all simultaneous fixed points is closed and is a proper class.*

(iv) *There exists a unique order-isomorphism* $J : \mathrm{On} \to K$ *enumerating the simultaneous fixed points of* $\langle F_i \mid i \in I \rangle$*. Moreover,* J *is a normal function.*

Proof. Define an ordinal function $F : \mathrm{On} \to \mathrm{On}$ as $F(\alpha) \coloneqq \sup\{F_i(\alpha) \mid i \in I\}$. This is well defined since we can use replacement to guarantee that $\{F_i(\alpha) \mid i \in I\}$ is a set. Notice that ξ is a fixed point of F if and only if it is a simultaneous fixed point of $\langle F_i \mid i \in I \rangle$ since

$$F(\xi) = \xi \iff (\forall i \in I) F_i(\xi) \leq \xi \iff (\forall i \in I) F_i(\xi) = \xi.$$

The last equivalence holds because all F_i grow at least as fast as the identity function. Note that this also implies that F itself grows at least as fast as the identity function. Moreover, F is continuous. Indeed, if λ is a limit ordinal, then

$$\begin{aligned} F(\lambda) &= \sup\{F_i(\lambda) \mid i \in I\} \\ &= \sup\{\sup\{F_i(\delta) \mid \delta < \lambda\} \mid i \in I\} && \dots \text{each } F_i \text{ is continuous} \\ &= \sup\{\sup\{F_i(\delta) \mid i \in I\} \mid \delta < \lambda\} \\ &= \sup\{F(\delta) \mid \delta < \lambda\}. \end{aligned}$$

In general, F is not normal, as it is not guaranteed to be increasing. However, notice that the proof of Theorem 3.12 only uses the fact that F is continuous and that it grows at least as fast as the identity. Hence, we can use the theorem on F to obtain (i), (ii) and (iii). Corollary 3.13 then gives us (iv). □

3.2 Ordinal Arithmetic

At last, we will define operations such as ordinal addition and multiplication. Before proceeding further, I highly recommend watching the video [26] by Vsauce, which illustrates the concepts of constructing larger ordinals from earlier ones in a very illustrative and intuitive way.

3.2.1 Definitions and Intuition

Definition 3.15. Let α and β be ordinals. We define ordinal numbers

(a) $\alpha + \beta$ as the order type of the set $(\{0\} \times \alpha) \cup (\{1\} \times \beta)$ when ordered lexicographically,

(b) $\alpha \cdot \beta$ as the order type of the set $\beta \times \alpha$ when ordered lexicographically.

Using the popular "matchstick" representation of ordinals used in [26], $\alpha + \beta$ can be imagined as a pile of decreasing matchsticks labeled by α, followed by another pile of matchsticks labeled by β. Notice that our previous notation is consistent, as $\alpha + 1 = \alpha \cup \{\alpha\} = \alpha^+$. We first use the elements of α to label the first pile, and we need one additional ordinal to label the second pile (which contains only a single matchstick).

Notice that we are using $\beta \times \alpha$ in the definition of $\alpha \cdot \beta$. The ordinal $\alpha \cdot \beta$ can be imagined as taking multiple piles of matchsticks labeled by α and arranging them next to each other. How should the piles be arranged? In a way that we need β to label them.

With this intuition, it should not be surprising that ordinal addition and multiplication are generally not commutative. It is easy to see that $1 + \omega = \omega$ (label the first pile by 0 and the other pile by $\omega \setminus \{0\}$), but $\omega + 1 \neq \omega$. For multiplication, consider $2 \cdot \omega$, the order type of countably infinitely many copies of $\{0, 1\}$ stacked behind each other. This can be clearly labeled by ω, so $2 \cdot \omega = \omega$. But $\omega \cdot 2$ is the order type of two consecutive copies of ω. When we try to label them using ω, we use all $n \in \omega$ to label the first copy and need more ordinals for the second copy. Therefore $\omega \cdot 2 > \omega$.

This may seem troubling, as commutativity is a fundamental property of arithmetic on the natural numbers. Nevertheless, we will soon see that this familiar behavior is retained.

Observation 3.16. *For any ordinals α, β, γ and natural $n \in \omega$ it holds that*

(a) $\alpha + 0 = \alpha = 0 + \alpha, \qquad \alpha \cdot 0 = 0 = 0 \cdot \alpha, \qquad \alpha \cdot 1 = \alpha = 1 \cdot \alpha,$

(b) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma, \qquad \alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma,$

(c) $\alpha \cdot 2 = \alpha + \alpha, \qquad \alpha \cdot 3 = \alpha + \alpha + \alpha, \qquad \alpha \cdot (n + 1) = \alpha \cdot n + \alpha.$

Definition 3.17. For ordinal numbers α and β, we define α^β recursively as

(i) $\alpha^0 \coloneqq 1$,

(ii) if $\beta = \gamma + 1$ is isolated, then $\alpha^\beta \coloneqq \alpha^\gamma \cdot \alpha$,

(iii) if β is a limit ordinal, then $\alpha^\beta \coloneqq \sup\{\alpha^\gamma \mid 0 < \gamma < \beta\}$.

To get an intuition for ordinal powers, consider the ordinal $\omega^2 = \omega \cdot \omega$. It represents multiple copies of ω arranged in the same manner as ω.

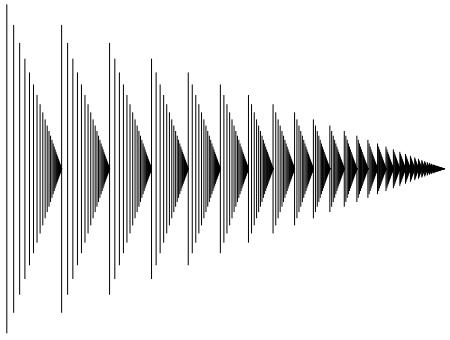

Figure 2: A representation of the ordinal ω^2. Each stick corresponds to an ordinal of the form $\omega \cdot m + n$ where m and n are natural numbers; [14].

To construct $\omega^3 = (\omega \cdot \omega) \cdot \omega$, we take multiple copies of ω^2 and arrange them in a way that requires ω to label them. If we repeat this process countably infinitely many times, we arrive at ω^ω.

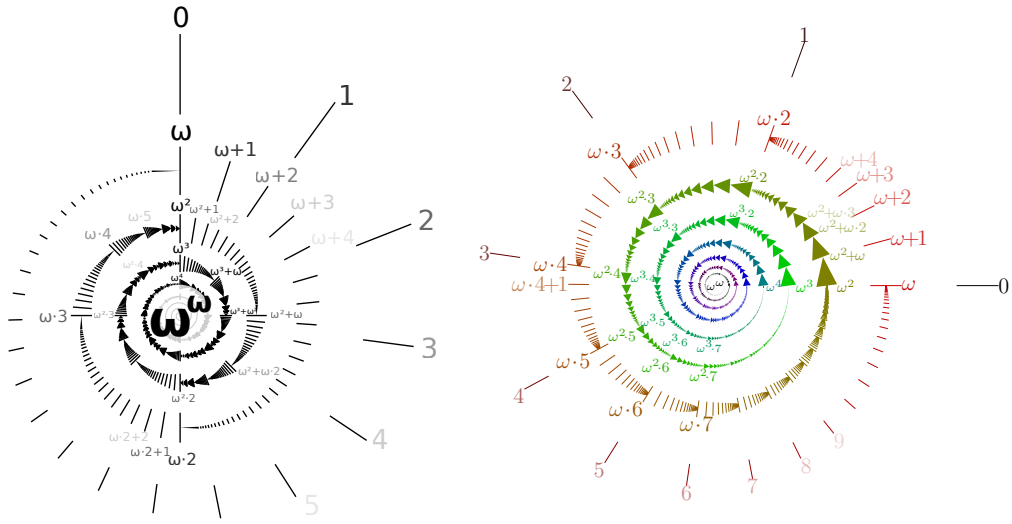

Figure 3: A spiral representation of ordinals up to ω^ω. One full turn corresponds to the mapping $f(\alpha) = \omega \cdot (1 + \alpha)$. Since ω^ω is the smallest fixed point of f, larger ordinals cannot be represented in this way; [6] and [4].

We can continue and arrive at larger and larger ordinals, such as $\omega^{(\omega^\omega)}$ or $\omega^{(\omega^{(\omega^\omega)})}$. That is a lot of parentheses, so from now on, we will write ω^{ω^ω} instead of $\omega^{(\omega^\omega)}$ and use parentheses only when we mean to say $(\omega^\omega)^\omega$.

Observation 3.18. *For any ordinals α and $\beta > 0$, it holds that*

(a) $0^0 = 1, \quad 0^\beta = 1,$

(b) $1^0 = 1, \quad 1^\beta = 1,$

(c) $\alpha^0 = 1, \quad \alpha^1 = \alpha, \quad \alpha^2 = \alpha \cdot \alpha, \quad \alpha^3 = (\alpha \cdot \alpha) \cdot \alpha.$

3.2.2 Basic Properties of Ordinal Operations

You should now have an intuition for how ordinal numbers constructed using these standard operations look. We continue by proving some of their basic properties.

Lemma 3.19 (Monotonicity of sum)**.** *For any ordinals α and β, it holds that*

(a) $\alpha < \beta \implies \gamma + \alpha < \gamma + \beta,$

(b) $\alpha < \beta \implies \alpha + \gamma \leq \beta + \gamma.$

Proof. (a) From the definition of addition and order types, it is easy to see that $\gamma + \alpha$ is an initial segment of $\gamma + \beta$. (b) The set of ordered pairs that defines $\alpha + \gamma$ is a subset of the set of ordered pairs that defines $\beta + \gamma$. And Lemma 3.3 states that the order type of the first is at most that of the second. □

Lemma 3.20 (Monotonicity of product)**.** *For any ordinals α, β and $\gamma > 0$, it holds that*

(a) $\alpha < \beta \implies \gamma \cdot \alpha < \gamma \cdot \beta$,

(b) $\alpha < \beta \implies \alpha \cdot \gamma \leq \beta \cdot \gamma$. *...for $\gamma = 0$ also holds*

Proof. (a) If $\alpha < \beta$, then $\alpha \times \gamma$ is an initial segment of $\beta \times \gamma$ when ordered lexicographically. (b) If $\alpha < \beta$, then $\gamma \times \alpha \subseteq \gamma \times \beta$, and the claim follows from Lemma 3.3. □

Remark. Note that the second statement in the two preceding lemmas does not, in general, hold under strict inequality. For example, $1 < 2$, but

$$1 + \omega = 2 + \omega = \omega, \qquad \text{and} \qquad 1 \cdot \omega = 2 \cdot \omega = \omega.$$

In fact, for any natural $n \in \omega$, we have that $n + \omega = \omega$ and $n \cdot \omega = \omega$.

Lemma 3.21 (Distributivity)**.** *For any ordinals α and β_1, β_2, we have*

$$\alpha \cdot (\beta_1 + \beta_2) = \alpha \cdot \beta_1 + \alpha \cdot \beta_2.$$

That is, ordinal addition and multiplication are left-distributive. However, in general, they are not right-distributive. Meaning that for some α and β_1, β_2

$$(\beta_1 + \beta_2) \cdot \alpha \neq \beta_1 \cdot \alpha + \beta_2 \cdot \alpha$$

Proof. Left distributivity essentially states that if we arrange $\beta + \gamma$ copies of α next to each other, then it is the same as first arranging β copies of α, followed by γ copies of α. This is obviously true from how we defined addition and multiplication. However, in general, these operations are not right-distributive. Consider $(1+1)\cdot\omega \neq \omega+\omega$. On the left, we have $2\cdot\omega$, and on the right, $\omega\cdot 2$. □

Theorem 3.22. *If m, n and k are natural numbers, then $m + n$, $m \cdot n$, and m^n are also natural numbers. Furthermore*

$$m + n = n + m, \qquad m \cdot n = n \cdot m, \qquad (m + n) \cdot k = m \cdot k + n \cdot k.$$

That is: addition and multiplication of natural numbers is commutative and right-distributive.

Proof. It is easy to see that $m + n$ and $m \cdot n$ are finite ordinals, and one can use induction on n to show that m^n is also finite.

To show that $m \cdot n = n \cdot m$, construct a bijection $f : n \times m \to m \times n$. We can use f and the lexicographic order on $n \times m$ to define a linear order $<_f$ on $m \times n$. It is well known (and it is proved in the basic set theory course) that any two linear orders on a finite set are isomorphic. Meaning that the two lexicographically ordered sets $n \times m$ and $m \times n$ are order-isomorphic; therefore, they have the same order types. Hence $m \cdot n = n \cdot m$. One can similarly show that addition commutes as well.

Right-distributivity is implied by commutativity and left-distributivity. □

Lemma 3.23 (Existence of the "right" difference)**.** *If $\alpha \leq \beta$, then there is a unique ordinal ϱ such that $\alpha + \varrho = \beta$. We denote ϱ by $\beta - \alpha$.*

Intuition. Any ordinal can be extenced by a specific amount to reach any larger ordinal.

Proof. If $\alpha \leq \beta$, then $\alpha = (\leftarrow, \alpha)$ is an initial segment of β, and its complement, $\beta \setminus \alpha$, is what we might denote as $[\alpha, \rightarrow)$. If ϱ is the order type of $\beta \setminus \alpha$, then clearly $\alpha + \varrho = \beta$. The uniqueness of ϱ follows from Lemma 3.19 (a). Suppose there were $\varrho_1 < \varrho_2$ satisfying $\alpha + \varrho_1 = \beta = \alpha + \varrho_2$. But since $\varrho_1 < \varrho_2$, we have that $\alpha + \varrho_1 < \alpha + \varrho_2$. □

Lemma 3.24 (Division with remainder)**.** *If $\beta > 0$, then for every ordinal α there are unique ordinals $\delta \leq \alpha$ and $\varrho < \beta$ such that $\alpha = \beta \cdot \delta + \varrho$.*

Intuition. Any ordinal α can be created by arranging multiple coppies of β in a specific way, and following this with a short tail ϱ.

Proof. Since $1 \leq \beta$, we have $\alpha \leq \beta \cdot \alpha$. If $\alpha = \beta \cdot \alpha$ (for example $\omega = 3 \cdot \omega$), choose $\delta \coloneqq \alpha$ and $\varrho \coloneqq 0$. It is not hard to show that the monotonicity of sum and product, together with left-distributivity, implies uniqueness; we will skip it.

If $\alpha < \beta \cdot \alpha$, let j be the isomorphism of the lexicographically ordered set $\alpha \times \beta$ and the ordinal $\beta \cdot \alpha$. Let $(\delta, \varrho) \in \alpha \times \beta$ be the (unique) pair mapped by j onto α. Necessarily $\delta < \alpha$ and $\varrho < \beta$. Since $\alpha \times \beta$ is ordered lexicographically, it is easy to see that $\alpha = \beta \cdot \delta + \varrho$. □

Lemma 3.25 (Monotonicity of power)**.** *For any ordinals α, β, γ and $\rho > 1$, it holds that*

(a) $\alpha < \beta \implies \alpha^\gamma \leq \beta^\gamma$,

(b) $\alpha < \beta \implies \rho^\alpha < \rho^\beta$.

Proof. (a) Using transfinite induction on γ. If $\gamma = 0$, then $\alpha^\gamma = \beta^\gamma = 1$. If $\gamma = \delta + 1$ and $\alpha^\delta \leq \beta^\delta$, from the monotonicity of product we have that

$$\alpha^\gamma = \alpha^\delta \cdot \alpha \leq \beta^\delta \cdot \beta = \beta^\gamma.$$

If γ is a limit ordinal and for every $\delta < \gamma$ already $\alpha^\delta \leq \beta^\delta$, then also

$$\alpha^\gamma = \sup\{\alpha^\delta \mid 0 < \delta < \gamma\} \leq \sup\{\beta^\delta \mid 0 < \delta < \gamma\} = \beta^\gamma.$$

(b) Suppose that $\rho > 1$. It is easy to show using transfinite induction on δ that for every $\delta > 1$ it holds that $\rho^\alpha < \rho^{\alpha+\delta}$. If $\alpha < \beta$, then according to Lemma 3.23 there is a unique $\delta > 0$ satisfying $\beta = \alpha + \delta$. □

Remark. Note that the first statement in the previous lemma does not, in general, hold under the strict inequality, even if $\gamma > 0$. For example, $2 < 3$, but $2^\omega = 3^\omega = \omega$. In general, if $n \in \omega$, then $n^\omega = \omega$.

Lemma 3.26 (Continuity in the second argument)**.** *The ordinal function*

(a) $F(\xi) = \alpha + \xi$ *is normal for every* $\alpha \geq 0$,

(b) $F(\xi) = \alpha \cdot \xi$ *is normal for every* $\alpha > 0$,

(c) $F(\xi) = \alpha^\xi$ *is normal for every* $\alpha > 1$.

Proof. All of the functions mentioned above are increasing for the specified α, since the respective operations are monotonic. We claim that they are also continuous. Let λ be a limit ordinal; we claim that $F(\lambda) = \sup\{F(\xi) \mid \xi < \lambda\}$. Suppose there were σ such that for all $\xi < \lambda$ we have $\sigma > F(\xi)$, but $\sigma < F(\lambda)$.

(a) $\alpha + \lambda \stackrel{?}{=} \sup\{\alpha + \xi \mid \xi < \lambda\}$: Lemma 3.23 claims that there is a unique ϱ such that $\sigma = \alpha + \varrho$. Because $\sigma < F(\lambda)$ we have that $\alpha + \varrho < \alpha + \lambda$ and thus $\varrho < \lambda$. It should hold that $\sigma > F(\varrho)$, but $\sigma = \alpha + \varrho = F(\varrho)$.

(b) $\alpha \cdot \lambda \stackrel{?}{=} \sup\{\alpha \cdot \xi \mid \xi < \lambda\}$: Lemma 3.24 claims the existence of unique ordinals δ and $\varrho < \alpha$ such that $\sigma = \alpha \cdot \delta + \varrho$. Since $\sigma < \alpha \cdot \lambda$ we have that $\delta < \lambda$ (monotonicity) and also $\delta + 1 < \lambda$ (λ is limit). Combining these we get

$$\sigma = \alpha \cdot \delta + \varrho < \alpha \cdot \delta + \alpha = \alpha \cdot (\delta + 1) = F(\delta + 1).$$

But we assumed that $\sigma > F(\xi)$ for all $\xi < \lambda$.

(c) $\alpha^\lambda \stackrel{?}{=} \sup\{\alpha^\xi \mid \xi < \lambda\}$: This just the definition of α^λ for $\alpha > 1$. □

Lemma 3.27 (Addition and multiplication in the exponent)**.** *For any ordinals α, β and γ, it holds that*

(a) $\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$,

(b) $(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$

Proof. (a) Trivially holds for $\alpha \leq 1$. Suppose $\alpha > 1$, we will use transfinite induction on γ. If $\gamma = 0$, there is nothing to prove. If $\gamma = \delta + 1$ is isolated, then $\beta + \gamma = (\beta + \delta) + 1$, and the statement follows from the induction hypothesis:

$$\alpha^{(\beta+\delta)+1} = \alpha^{(\beta+\delta)} \cdot \alpha^1 = \alpha^\beta \cdot \alpha^\delta \cdot \alpha = \alpha^\beta \cdot \alpha^{\delta+1} = \alpha^\beta \cdot \alpha^\gamma.$$

Finally, if γ is a limit ordinal and the statement holds for all $\gamma < \delta$, then

$$\begin{aligned}
\alpha^\beta \cdot \alpha^\gamma &= \sup\{\alpha^\beta \cdot \xi \mid \xi < \alpha^\gamma\} && \dots F(\xi) = \alpha^\beta \cdot \xi \text{ is normal} \\
&= \sup\{\alpha^\beta \cdot \alpha^\delta \mid 0 < \delta < \gamma\} && \dots \alpha^\gamma = \sup\{\alpha^\delta \mid 0 < \delta < \gamma\} \\
&= \sup\{\alpha^{\beta+\delta} \mid 0 < \delta < \gamma\} && \dots \text{induction hypothesis} \\
&= \sup\{\alpha^\varepsilon \mid 0 < \varepsilon < \beta + \gamma\} && \dots F(\xi) = \beta + \xi \text{ is normal} \\
&= \alpha^{\beta+\gamma}.
\end{aligned}$$

(b) Suppose that $\beta, \gamma \neq 0$ and $\alpha > 1$, otherwise, it trivially holds. We will again use transfinite induction on γ. If $\gamma = 0$, then it holds. If $\gamma = \delta + 1$ is isolated, then the statement follows from (a) and the induction hypothesis:

$$(\alpha^\beta)^{\delta+1} = (\alpha^\beta)^\delta \cdot (\alpha^\beta)^1 = \alpha^{\beta \cdot \delta} \cdot \alpha^\beta = \alpha^{\beta \cdot \delta + \beta} = \alpha^{\beta \cdot (\delta+1)} = \alpha^{\beta \cdot \gamma}.$$

Finally, if γ is a limit ordinal, then $\beta \cdot \gamma$ is also limit and

$$\begin{aligned}
(\alpha^\beta)^\gamma &= \sup\{(\alpha^\beta)^\delta \mid 0 < \delta < \gamma\} && \dots \gamma \text{ is limit} \\
&= \sup\{\alpha^{\beta \cdot \delta} \mid 0 < \delta < \gamma\} && \dots \text{induction hypothesis} \\
&= \sup\{\alpha^\varepsilon \mid 0 < \varepsilon < \beta \cdot \gamma\} && \dots F(\xi) = \beta \cdot \xi \text{ is normal} \\
&= \alpha^{\beta \cdot \gamma}.
\end{aligned}$$

The last equality holds because $\beta \cdot \gamma$ is a limit ordinal. □

3.2.3 Ordinal Equations and Power Expansions

Example. Suppose we want to find all ξ and β satisfying $\xi + \beta = \omega$. Lemma 3.23 claims that $\xi \leq \omega$ and $\beta = \omega - \xi$. Suppose $\xi = \omega$, then $\beta = 0$. If $\xi = n$ is a natural number, then $\beta = \omega - n = \omega$. We conclude that β can attain only two different values.

Proposition 3.28. *Let α be an ordinal and consider the equation $\xi + \beta = \alpha$. The set of solutions (ξ, β) contains only finitely many distinct values of β.*

Proof. Suppose that for some α, there are infinitely many distinct values of β in the solution set. Let $(\xi_n, \beta_n)_{n\in\omega}$ be a sequence of solutions such that $\beta_n < \beta_{n+1}$ for all n. Since $\xi_n + \beta_n = \xi_{n+1} + \beta_{n+1}$, from the monotonicity of sum we have that $\xi_n > \xi_{n+1}$ for all $n \in \omega$. We have constructed an infinite strictly decreasing sequence, which is impossible since On is well-ordered. □

We are able to express any natural number n as an expansion of powers of any base $b > 1$. We will prove that a similar statement holds for ordinal numbers too. A base of special importance is ω (as it is the first transfinite ordinal), and the expansion of α over ω is called its *Cantor normal form*; however, an expansion is possible over any base $\beta > 1$.

Lemma 3.29. *If $k, m_0, m_1, \ldots, m_k$ are natural numbers and $\gamma_0, \gamma_1, \ldots, \gamma_k > \delta$ are ordinals, then*

$$\omega^\delta > \omega^{\gamma_0} \cdot m_0 + \omega^{\gamma_1} \cdot m_1 + \cdots + \omega^{\gamma_k} \cdot m_k.$$

Proof. Let m be the largest among all m_i, and γ be the largest among all γ_i. Then $\omega^\gamma \cdot m \cdot k$ is an upper bound of the sum on the right side of the equation. We assumed that $\delta \geq \gamma + 1$, so $\omega^\delta \geq \omega^{\gamma+1} > \omega^\gamma \cdot m \cdot k$. □

Theorem 3.30 (Expansion over ω)**.** *For any $\alpha > 0$ there are unique natural numbers $k, m_0, m_1, \ldots, m_k \neq 0$ and ordinals $\gamma_0 > \gamma_1 > \cdots > \gamma_k$ which satisfy*

$$\alpha = \omega^{\gamma_0} \cdot m_0 + \omega^{\gamma_1} \cdot m_1 + \cdots + \omega^{\gamma_k} \cdot m_k. \tag{3.1}$$

The sum on the right side of the equation is called the Cantor normal form of α.
Furthermore, if

$$\beta = \omega^{\delta_0} \cdot n_0 + \omega^{\delta_1} \cdot n_1 + \cdots + \omega^{\delta_l} \cdot n_l \tag{3.2}$$

is the Cantor normal form of an ordinal β, then $\beta > \alpha$ if and only if one of the two following cases occurs:

(a) *$l > k$ and the first k terms of β are idential to those of α. For example: $\alpha = \omega^2 + \omega \cdot 2$ and $\beta = \omega^2 + \omega \cdot 2 + 3$.*

(b) *there exists an index $i \leq \min(k, l)$ at which (γ_i, m_i) and (δ_i, n_i) differ, and for the smallest such index i either $\delta_i > \gamma_i$, or $\delta_i = \gamma_i$ and $n_i > m_i$. For example $\alpha = \omega^2 + \omega \cdot 2$ and $\beta = \omega^2 \cdot 5 + \omega \cdot 2$.*

Proof. We prove the first part by transfinite induction on α. The CNF of $\alpha = 1$ is $\alpha = \omega^0 \cdot 1$. Suppose $\alpha > 0$ and that every nonzero $\beta < \alpha$ has a unique CNF. The ordinal function $\gamma \mapsto \omega^\gamma$ is normal, so according to Lemma 3.10, there exists a maximal ordinal γ such that $\omega^\gamma \leq \alpha$. Similarly, from the normality of product in the second argument follows the existence of a maximal ordinal δ such that $\omega^\gamma \cdot \delta \leq \alpha$. Also, $\delta < \omega$, since $\omega^{\gamma+1} = \omega^\gamma \cdot \omega > \alpha$, which contradicts the choice of γ. If $\omega^\gamma \cdot \delta = \alpha$, then the uniqueness of this expansion follows from Lemma 3.29. If $\omega^\gamma \cdot \delta < \alpha$, then there exists a unique ordinal $\beta = \alpha - \omega^\gamma \cdot \delta$ such that $\omega^\gamma \cdot \delta + \beta = \alpha$. Note that $\beta < \omega^\gamma$; otherwise, we get $\omega^\gamma \cdot \delta + \beta \geq \omega^\gamma \cdot (\delta + 1)$, which contradicts the choice of δ.

To find the CNF on α, let

$$\beta = \omega^{\gamma_1} \cdot m_1 + \omega^{\gamma_2} \cdot m_2 + \cdots + \omega^{\gamma_k} \cdot m_k$$

be the CNF of β. Define $\gamma_0 \coloneqq \gamma$ and $m_0 \coloneqq \delta$. Then $\gamma_0 > \gamma_1$, and (3.1) is the CNF of α. The uniqueness of this expansion follows from Lemma 3.29 and the unique choice of β.

Next, we prove the second part of the theorem. Suppose that the ordinals α and β have Cantor normal forms (3.1) and (3.2). If (a) holds, $\beta > \alpha$ because the trailing terms in the expansion of β are nonzero. Suppose that (b) holds and that i is the least index at which the two expansions differ. If $\delta_i > \gamma_i$, then $\beta > \alpha$ from Lemma 3.29. If $\delta_i = \gamma_i$ and $n_i > m_i$, then $n_i \geq m_i + 1$ and

$$\omega^{\delta_i} \cdot n_i \geq \omega^{\gamma_i} \cdot m_i + \omega^{\gamma_i}.$$

Lemma 3.29 claims that the second summand on the right (ω^{γ_i}) is a strict upper bound of the remaining summands in (3.1), the expansion of α; thus $\beta > \alpha$.

All that remains is to prove the reverse implication. If $\beta > \alpha$, then their Cantor normal forms (3.1) and (3.2) have to differ. Either the CNF of one of the ordinals is the same as the beginning of the CNF of the other, or there exists an index at which they differ. We can use the already proven implication to show that the only two possible cases are (a) and (b). □

Corollary 3.31 (Alternative expansions)**.** *For any $\alpha > 0$, it holds that*

(a) *there is a unique natural number $l > 0$ and unique ordinals $\gamma_0 \geq \gamma_1 \geq \cdots \geq \gamma_l$ which satisfy*

$$\alpha = \omega^{\gamma_0} + \omega^{\gamma_1} + \cdots + \omega^{\gamma_l},$$

(b) *there are unique ordinals β and γ such that*

$$\alpha = \omega^\gamma \cdot (\beta + 1).$$

Proof. (a) For any ordinal γ and natural m, is the ordinal number $\omega^\gamma \cdot m$ equal to the sum of m summands of the form ω^γ. We obtain the expansion in (a) by expressing each term in the CNF of α in this expanded form.

(b) If α has CNF (3.1), we let $\gamma = \gamma_k$. Then for all $i \leq k$ is $\gamma_i = \gamma + \delta_i$ for $\delta_i = \gamma_i - \gamma$. From the properties of exponents and left-distributivity, we get

$$\alpha = \omega^\gamma \cdot \big(\omega^{\delta_0} \cdot m_0 + \omega^{\delta_1} \cdot m_1 + \cdots + \omega^0 \cdot m_k\big).$$

The parentheses on the right contain an isolated ordinal $\beta + 1$, because m_k is a nonzero natural number. The uniqueness of the ordinals γ and β follows from the uniqueness of the CNF of α. □

Theorem 3.32 (Expansion over any base)**.** *The choice of ω as a base in Theorem 3.30 was arbitrary; the same holds for any ordinal base $\beta > 1$. We just need to restrict the nonzero coefficients $m_0, m_1, \ldots, m_k$ to be smaller than β.*

Proof. We did not use any special properties of ω in the proof of Theorem 3.30, so we only need to modify Lemma 3.29. If we slightly change its claim, only for decreasing exponents $\gamma_0 > \gamma_1 > \cdots > \gamma_k$ and coefficients $m_i < \beta$, we can prove it by transfinite induction on γ_i. □

Remark. If we restrict ourselves only to natural numbers, we obtain the familiar theorem about expanding natural numbers using powers of a base $b > 1$.

3.3 Countable and Uncountable Ordinals

We saw earlier that for all natural numbers n, it holds that $n+\omega = n\cdot\omega = n^\omega = \omega$. We also proved that the functions corresponding to these basic operations,

$$A_n(\xi) = n + \xi, \qquad M_n(\xi) = n \cdot \xi, \qquad E_n(\xi) = n^\xi,$$

are normal. Theorem 3.12 claims that each of them has infinitely many fixed points. It is easy to see that no (nonzero) natural number is a fixed point, and above we have observed that ω is a fixed point of all of them. It is, in fact, the smallest (nonzero) fixed point. Notice that this makes intuitive sense. When restricted to natural numbers, these are all fast growing functions ($A \ll M \ll E$), so we need a new concept (countable infinity) to find a fixed point.

Now consider what would happen if we replaced n with ω and tried to find a (nonzero) fixed point of these new ω-functions. Theorem 3.12 claims that the smallest such fixed points are:

- $F_A = \sup\{0, \omega, \omega+\omega, \omega+\omega+\omega, \omega\cdot 4, \omega\cdot 5, \ldots\} = \omega\cdot\omega$,
- $F_M = \sup\left\{1, \omega, \omega\cdot\omega, \omega\cdot\omega\cdot\omega, \omega^4, \omega^5, \ldots\right\} = \omega^\omega$,
- $F_E = \sup\left\{1, \omega, \omega^\omega, \omega^{\omega^\omega}, \omega^{\omega^{\omega^\omega}}, \ldots\right\}$, and we no longer have notation to describe this number; we will denote it as ε_0.

The question is: did we escape the countable infinity represented by ω? No, we will soon see that all of these numbers are, in fact, still countable. Nonetheless, we have stumbled upon something important. The last number, $\varepsilon_0 = \omega^{\varepsilon_0}$, is closely connected to Peano arithmetic, and we will also use it when proving Goodstein's theorem. It also gives rise to an entire class of ordinals called the *epsilon numbers*.

3.3.1 Epsilon Numbers

Definition 3.33. An ordinal ξ is an *epsilon number* if it is a fixed point of the normal function $\xi \mapsto \omega^\xi$. That is, if $\xi = \omega^\xi$. Corollary 3.13 asserts the existence of a normal function $\varepsilon : \mathrm{On} \to \{\xi \mid \xi = \omega^\xi\}$; we call it the *epsilon function* and denote the ordinal $\varepsilon(\beta)$ as ε_β.

Proposition 3.34. *For any ordinal β, it holds that*

(i) $\varepsilon_0 = \sup\left\{1, \omega, \omega^\omega, \omega^{\omega^\omega}, \omega^{\omega^{\omega^\omega}}, \ldots\right\}$,

(ii) $\varepsilon_{\beta+1} = \sup\left\{1, \varepsilon_\beta, \varepsilon_\beta{}^{\varepsilon_\beta}, \varepsilon_\beta{}^{\varepsilon_\beta{}^{\varepsilon_\beta}}, \varepsilon_\beta{}^{\varepsilon_\beta{}^{\varepsilon_\beta{}^{\varepsilon_\beta}}}, \dots\right\}$,

(iii) $\varepsilon_\beta = \sup\{\varepsilon_\delta \mid \delta < \beta\}$, *whenever* β *is a limit ordinal.*

Proof. We prove the theorem by transfinite induction on β. (i) This is the definition of ε_0. (iii) Holds because the epsilon function is normal. (ii) Following Theorem 3.12, we know that $\varepsilon_{\beta+1}$ is the limit of the sequence

$$\varepsilon_\beta + 1, \omega^{\varepsilon_\beta+1}, \omega^{\omega^{\varepsilon_\beta+1}}, \omega^{\omega^{\omega^{\varepsilon_\beta+1}}}, \dots$$

Let β_n denote the element with index $n \in \omega$. Define a different sequence for $n \geq 2$ as $\beta'_2 \coloneqq \varepsilon_\beta{}^\omega$ and $\beta'_{n+1} \coloneqq \varepsilon_\beta{}^{\beta'_n}$. Clearly

$$\sup\{\beta'_n \mid n \geq 2\} = \sup\left\{1, \varepsilon_\beta, \varepsilon_\beta{}^{\varepsilon_\beta}, \varepsilon_\beta{}^{\varepsilon_\beta{}^{\varepsilon_\beta}}, \varepsilon_\beta{}^{\varepsilon_\beta{}^{\varepsilon_\beta{}^{\varepsilon_\beta}}}, \dots\right\}.$$

We will use induction on n to show $\beta_n = \beta'_n$ for all $n \geq 2$:

$$\begin{aligned}
\beta_1 &= \omega^{\varepsilon_\beta+1} = \omega^{\varepsilon_\beta} \cdot \omega = \varepsilon_\beta \cdot \omega \\
\beta_2 &= \omega^{\omega^{\varepsilon_\beta+1}} = \omega^{(\varepsilon_\beta \cdot \omega)} = (\omega^{\varepsilon_\beta})^\omega = \varepsilon_\beta{}^\omega = \beta'_2 \\
\beta_3 &= \omega^{\omega^{\omega^{\varepsilon_\beta+1}}} = \omega^{\varepsilon_\beta{}^\omega} = \omega^{\varepsilon_\beta{}^{1+\omega}} = \omega^{\varepsilon_\beta \cdot \varepsilon_\beta{}^\omega} = (\omega^{\varepsilon_\beta})^{\varepsilon_\beta{}^\omega} = \varepsilon_\beta{}^{\varepsilon_\beta{}^\omega} = \beta'_3 \\
\beta_{n+2} &= \omega^{\beta_{n+1}} = \omega^{\beta'_{n+1}} = \omega^{\varepsilon_\beta{}^{\beta'_n}} = \omega^{\varepsilon_\beta{}^{1+\beta'_n}} = \omega^{\varepsilon_\beta \cdot \varepsilon_\beta{}^{\beta'_n}} = (\omega^{\varepsilon_\beta})^{\beta'_{n+1}} = \varepsilon_\beta{}^{\beta'_{n+1}} = \beta'_{n+2}
\end{aligned}$$ □

Lemma 3.35 (AC_ω)**.** *A countable union of countable sets is countable. Specifically, if* β *and* γ_α *for* $\alpha < \beta$ *are countable ordinals, then* $\gamma = \sup\{\gamma_\alpha \mid \alpha < \beta\}$ *is also a countable ordinal.*

Remark. AC_ω denotes the *axiom of countable choice*, which states that every countable set has a choice function.

Proof. Let $A = \langle A_n \mid n \in I \rangle$ be a countable collection of sets, WLOG $I = \omega$, such that all A_n are countable, and denote $S \coloneqq \bigcup A$. We will define an injection $g : S \to \omega \times \omega$ (here, $\omega \times \omega$ is countable[6]). Since each A_n is countable, it injects into ω, and we can *choose* an injection $j_n : A_n \to \omega$ (and because there are only countably many sets A_n, we are making only countably many choices). For an element $a \in S$, define

$$n_a \coloneqq \min\{n \in \omega \mid a \in A_n\}.$$

This number indicates in which A_n does a first appear in. Notice that more elements $a \in S$ can have the same number n_a, but that $j_{n_a}(a)$ uniquely identifies a among these elements (since j_{n_a} is injective). Hence, we can define an injection $g : a \mapsto (n_a, j_{n_a}(a))$. □

From now on, we will generally assume AC_ω.

Lemma 3.36 (AC_ω)**.** *The ordinal* ω^α *is countable* $\iff$ α *is countable.*

[6]Prove that the Cartesian product of finitely many countable sets is countable. Hint: prime numbers might help. Then try proving it without the use of primes.

Proof. We first prove '$\Leftarrow$' by transfinite induction on α. If $\alpha = 0$, it holds. Now suppose the claim holds for a countable α and consider the ordinal $\omega^{\alpha+1} = \omega^\alpha \cdot \omega$. This is countable because it is the order type of the set $\omega \times \omega^\alpha$, which is countable since it is the cartesian product of two countable sets. Finally, if α is a countable limit ordinal, then

$$\omega^\alpha = \sup\{\omega^\delta \mid \delta < \alpha\} = \bigcup\{\omega^\delta \mid \delta < \alpha\}.$$

Because α is countable and all ω^δ are countable ordinals (induction hypothesis), Lemma 3.35 implies that ω^α is countable as well.

We prove '$\Rightarrow$' by contraposition. Suppose that α is uncountable; since $\xi \mapsto \omega^\xi$ is a normal function, it grows at least as fast as the identity function, and so $\omega^\alpha \geq \alpha$. We can now easily define an injection $\alpha \to \omega^\alpha$, showing that ω^α is uncountable. □

Lemma 3.37. *The epsilon number ε_0 is countable.*

Proof. By definition, ε_0 is the limit of the sequence $\alpha_0 = 1$ and $\alpha_{n+1} = \omega^{\alpha_n}$. By induction on n and using the previous lemma, one can show that all α_n are, in fact, countable. This implies that ε_0 is a countable union of countable ordinals and is therefore countable as well. □

Note on ordinal notations We have proven the previous lemma via a statement that requires AC_ω, because it made our job easier (and will continue to do so greatly down the line). However, we could have shown that ε_0 is countable by realizing that every $\alpha < \varepsilon_0$ has a finite *hereditary* Cantor normal form; therefore, we can use prime numbers to encode α as a unique natural number, hence constructing an injection $\varepsilon_0 \to \omega$. Hereditary CNF simply means that if any of the powers γ_i are ordinals larger than ω, then we express them in CNF as well, and we repeat the process inductively. For example:

$$\alpha = \omega^{\omega^{\omega+1}+\omega^2\cdot 3+5} + \omega^{\omega\cdot 2+1} + \omega \cdot 2 + 7.$$

But ε_0 cannot be represented by a finite hereditary CNF, since $\varepsilon_0 = \omega^{\varepsilon_0}$ is its CNF, which is self-referential.

This concept can be generalized: if λ is a large ordinal, and every $\alpha < \lambda$ can be represented as a finite syntactic structure over some finite alphabet, then we can encode α as a unique natural number, constructing an injection $\lambda \to \omega$. This concept is called *ordinal notations*, and it does not require any kind of choice. However, the larger the ordinal λ, the more complicated the finite structure and encoding become. It is crucial to understand that when we use AC_ω to prove the countability of a large ordinal (like ε_0 or Γ_0 later), we are using it to simplify the process, not because we *have to* use it.

However, general statements like Lemma 3.36 or the following theorem do require the use of countable choice and cannot be proved in bare ZF.

Theorem 3.38 (AC_ω)**.** *The epsilon number ε_β is countable $\iff$ β is countable.*

Proof. We first prove '$\Leftarrow$' by transfinite induction on β. The base case $\beta = 0$ has been verified by the previous lemma. Suppose the claim holds for a countable

ordinal β; that is, ε_β is countable, and we want to show that $\varepsilon_{\beta+1}$ is countable as well. Theorem 3.12 claims that $\varepsilon_{\beta+1}$ is the limit of the sequence $\alpha_0 = \varepsilon_\beta + 1$ and $\alpha_{n+1} = \omega^{\alpha_n}$. Note that α_0 is countable since ε_β is countable. One can now use induction on n and Lemma 3.36 to show that all α_n are in fact countable. This implies that $\varepsilon_{\beta+1}$ is a countable union of countable ordinals and is therefore also countable.

Finally, assume that β is a countable limit ordinal. Because the epsilon function $\beta \mapsto \varepsilon_\beta$ is normal, it holds that

$$\varepsilon_\beta = \sup\{\varepsilon_\delta \mid \delta < \beta\} = \bigcup\{\varepsilon_\delta \mid \delta < \beta\}.$$

Since β is countable, this is a countable union of countable ordinals (induction hypothesis), so it is countable as well.

One can prove '$\Rightarrow$' in the same manner as in Lemma 3.36. □

The theorem we have just proven places us in a difficult position. Does an uncountable ordinal even exist? If we assume the axiom of choice, then it is fairly easy to find one: just well-order the uncountable set $\mathcal{P}(\omega)$ and take its order type. Finding one in $\mathsf{ZF} + \mathsf{AC}_\omega$ seems to be much more difficult.

3.3.2 The Veblen Hierarchy

We know that combining countable ordinals using the standard operations defined above produces more countable ordinals. The best tool for constructing large ordinals we currently have are the epsilon numbers (and, moreover, Theorem 3.12 implies that for any ordinal, there is a larger epsilon number), but it seems like they will not help us either. Consider the sequence

$$\gamma_0 = \varepsilon_0,\ \gamma_{n+1} = \varepsilon_{\gamma_n} \qquad \longrightarrow \qquad \varepsilon_0, \varepsilon_{\varepsilon_0}, \varepsilon_{\varepsilon_{\varepsilon_0}}, \varepsilon_{\varepsilon_{\varepsilon_{\varepsilon_0}}}, \dots$$

The largest number we can currently construct is the limit of this sequence; the first fixed point of the epsilon function $\xi \mapsto \varepsilon_\xi$, a number denoted as ζ_0. However, this number, while enormously large, is still countable. The reason is that all the terms γ_n are countable (by induction and the previous theorem); hence, ζ_0 is a countable union of countable ordinals and is thus also countable.

We could define *zeta* (ζ) *numbers* in a similar fashion to how we defined epsilon numbers; however, for the same reasons that epsilon numbers with countable indices are countable, we would arrive at the conclusion that any zeta number with a countable index is still countable.

We could even create an entire hierarchy of these special fixed-point numbers. The bottom tier would be $\varphi_0(\beta) = \omega^\beta$; the second tier would be the epsilon numbers $\varphi_1(\beta) = \varepsilon_\beta$; the third tier would be the zeta numbers $\varphi_2(\beta) = \zeta_\beta$; the next one would be the so-called *eta numbers* $\varphi_3(\beta) = \eta_\beta$, and so on. The tiers are defined in such a way that the values of φ_{n+1} are the fixed points of φ_n. We could now consider the ordinal

$$\Lambda \coloneqq \sup\{\varphi_n(0) \mid n < \omega\}.$$

However, this ordinal is *still countable*, as it is a countable union of countable ordinals.

The functions φ_n we have just constructed are called the *Veblen functions*, and they can be generalized for arbitrary ordinal indices.

Definition 3.39 (Veblen hierarchy, 1908)**.** The functions $\varphi_\alpha : \mathrm{On} \to \mathrm{On}$ are defined for all ordinals $\alpha \geq 0$ recursively as:

(i) $\varphi_0(\beta) = \omega^\beta$,

(ii) $\varphi_{\alpha+1}(\beta)$ is the βth fixed point of φ_α, starting at $\beta = 0$.

(iii) when α is a limit ordinal, we define $\varphi_\alpha(\beta)$ as the βth simultaneous fixed point of all the φ_δ for $\delta < \alpha$, also starting at $\beta = 0$.

Observation 3.40. *The function φ_α is normal for every α.*

Proof. By transfinite induction on α. It holds for $\alpha = 0$ because exponentiation is normal in the second argument. If $\alpha = \gamma + 1$ is isolated, then φ_α enumerates the fixed points of the normal function φ_γ and is by Corollary 3.13 normal. If α is isolated, it enumerates the simultaneous fixed points of the normal functions φ_δ for $\delta < \alpha$ and so φ_α is normal by Theorem 3.14 (iv). □

Exercise 8. Show that if $\gamma > \alpha$, then $\varphi_\alpha(\varphi_\gamma(\beta)) = \varphi_\gamma(\beta)$ for any β.

This demonstrates that the values of $\varphi_{\alpha+1}$ are not only fixed points of φ_α, but they are fixed points of all φ_δ for $\delta \leq \alpha$. This means that we could have used condition (iii) to define φ_α for all ordinals α, not only for limit ordinals.

Exercise 9. Show that the ordinal function $\alpha \mapsto \varphi_\alpha(0)$ is normal.

Definition 3.41. An ordinal worth noting is the *Feferman–Schütte ordinal* Γ_0, defined as the first fixed point of the function $\alpha \mapsto \varphi_\alpha(0)$, or equivalently, as the limit of the sequence $\gamma_0 = \varphi_0(0)$, $\gamma_{n+1} = \varphi_{\gamma_n}(0)$, that is

$$\gamma_0 = 1, \gamma_1 = \varphi_1(0) = \varepsilon_0, \gamma_2 = \varphi_{\varepsilon_0}(0) = \varphi_{\varphi_{\varphi_0(0)}(0)}(0), \gamma_3 = \varphi_{\varphi_{\varphi_{\varphi_0(0)}(0)}(0)}(0), \dots$$

It is one of the most famous ordinals in logic, and we will attempt to provide an explanation of why in Section 3.4.5.

Exercise 10. Show that Γ_0 is the first ordinal $\gamma > 0$ closed under φ. That is, the least $\gamma > 0$ such that for all $\alpha, \beta < \gamma$ we have $\varphi_\alpha(\beta) < \gamma$.

In other words, Γ_0 is the first ordinal that cannot be reached from below via repeated application of the Veblen functions. Does that mean that we have finally found an uncountable ordinal?

Exercise 11. Show that for all ordinals $\alpha < \Gamma_0$ it holds that $\varphi_\alpha(\Gamma_0) = \Gamma_0$.

That is, all φ_α attain the value Γ_0 at the same time! Even the very "slow" growing $\varphi_0(\beta) = \omega^\beta$ catches up to functions like $\varphi_{\varepsilon_0}(\beta)$ or $\varphi_{\varphi_{\varepsilon_0}(0)}(\beta)$, and they all momentarily synchronize at $\beta = \Gamma_0$.

Theorem 3.42 (AC_ω)**.** *The ordinal $\varphi_\alpha(\beta)$ is countable $\iff$ α, β are countable.*

Proof. For the '$\Leftarrow$' direction, we use transfinite induction on α. Notice that Lemma 3.36 is our base case (when $\alpha = 0$). If $\alpha = \gamma + 1$ is isolated, the proof is essentially the same as the proof of Theorem 3.38. If α is a countable limit ordinal, then we prove the claim by transfinite induction on β.

- If $\beta = 0$, then $\varphi_\alpha(0)$ is by Theorem 3.14 (ii) equal to the limit of the sequence

$$\gamma_0 = 0, \qquad \gamma_{n+1} = \sup\{\varphi_\delta(\gamma_n) \mid \delta < \alpha\}.$$

  Using our outer induction hypothesis (for α), one can show by induction on n that all γ_n are countable. Thus $\varphi_\alpha(0)$ is a countable union of countable ordinals and is therefore also countable.

- If $\beta = \beta' + 1$ is isolated, then the proof is almost identical to the case when $\beta = 0$. One just starts with $\gamma_0 = \varphi_\alpha(\beta') + 1$.

- If β is a limit ordinal, then since φ_α is normal we have that $\varphi_\alpha(\beta) = \sup\{\varphi_\alpha(\delta) \mid \delta < \beta\}$. Because we assume that β is countable, $\varphi_\alpha(\beta)$ is thus a countable union of countable ordinals (inner induction hypothesis) and is countable as well.

For the '$\Rightarrow$' direction, we need to show that $\varphi_\alpha(\beta) \geq \alpha, \beta$; so if either α or β is uncountable, then $\varphi_\alpha(\beta)$ is also uncountable. Clearly $\varphi_\alpha(\beta) \geq \beta$ because φ_α is a normal function. Exercise 9 implies that $\varphi_\alpha(0) \geq \alpha$, and from here we have that $\varphi_\alpha(\beta) \geq \varphi_\alpha(0) \geq \alpha$. □

Corollary 3.43. *The Feferman–Schütte ordinal Γ_0 is countable, as it is a countable union of countable ordinals.*

Remark. As mentioned in Section 3.3.1, the axiom of countable choice is not required to show that Γ_0 is countable, but it makes the task easier. If we wanted to prove it in ZF, we would show that the Veblen functions $\varphi_\alpha(\beta)$ provide a way of expressing every $\gamma < \Gamma_0$ as a unique finite syntactic structure. For more details about the Veblen hierarchy, see Sections 6.5, 7, and 8 of [10].

What we have shown in this section demonstrates an important concept in set theory when assuming AC_ω: one cannot reach uncountable infinity by starting from ω and applying ordinal operations such as addition, multiplication, exponentiation, finding fixed points of normal functions, and taking suprema — all the while utilizing only the ordinals we have already constructed along the way.

3.3.3 Hartogs' Theorem

Does that mean that all hope is lost and there are no uncountable ordinals? Thankfully, no. The following theorem gives us a way out.

Theorem 3.44 (Hartogs, 1915)**.** *For any set x, there exists an ordinal η such that there is no injection $\eta \to x$. The least such η is called the* Hartogs number *of x.*

Proof (cf. [9]*).* Consider the set (why is this a set?)

$$\mathcal{W} = \{(A, <_R) \mid A \subseteq x \text{ and } <_R \text{ is a well-ordering of } A\}.$$

We can use replacement to construct the set

$$S = \{\alpha \in \mathrm{On} \mid \text{there exists } (A, <_R) \in \mathcal{W} \text{ order-isomorphic to } \alpha\}$$

by assigning to each $(A, <_R)$ its order type.

But this set is exactly the Hartogs number of x. Notice that S is transitive: if $\alpha \in S$ and $\gamma < \alpha$, then $\alpha \in S$ as well. A transitive set of ordinals is again an ordinal, so S is an ordinal number η. Furthermore, there is no injection from η into x, because if there were, then we would get the contradiction that $\eta \in \eta$. And finally, η is the least such ordinal. If $\alpha < \eta$, then also $\alpha \in \eta$, and there is an injection $\alpha \to x$. □

Remark. It is crucial to note that the theorem we just proved, which gives us the Hartogs number as a von Neumann ordinal, is more powerful than Hartogs' original 1915 result. Hartogs, working in Z (proposed in 1908 by Zermelo, containing the axioms of ZF except replacement and foundation), only proved the existence of *a* well-ordered set that could not be injected into x; but he did not—and could not—show it was a von Neumann ordinal. The general theorem that "every well-ordered set is isomorphic to a unique von Neumann ordinal" is itself not provable in Z and requires replacement (see proof of Theorem 2.13). We will now use this modern, replacement-based construction to construct an uncountable ordinal as the Hartogs number of ω. It is this very step, guaranteeing that the collection of all countable ordinals is a set, that fails in Z and was one of the motivations for Fraenkel and Skolem to propose the axiom of replacement in 1922.

Notice that the theorem does not say that $x \prec \eta$, because this does not in general hold without AC. However, if x is well-ordered, then it has an order type α, and we can compare α with η.

This allows us to access an uncountable ordinal. Let ω_1 be the Hartogs number of ω. Either $\omega_1 \leq \omega$, which is impossible since there would be an injection $\omega_1 \to \omega$, or $\omega_1 > \omega$, which gives us a way to construct an injection $\omega \to \omega_1$; therefore, $\omega \preceq \omega_1$. But since $\omega_1 \not\preceq \omega$, the Cantor–Berstein theorem implies that $\omega \prec \omega_1$. By the definition of Hartogs numbers, ω_1 is the least ordinal with this property, meaning that it is the *first uncountable ordinal*,[7] and we can write

$$\omega_1 = \{\alpha \in \mathrm{On} \mid \alpha \preceq \omega\}$$

It is almost impossible to grasp just how unfathomably large ω_1 is. The entire vast, complex, mind-boggling hierarchy of ordinals described by the Veblen functions up to Γ_0 is still just a tiny, countable speck at the absolute "bottom" of the ordinal line from the perspective of ω_1.

Exercise 12. Show that for any countable ordinal α, it holds that $\varphi_\alpha(\omega_1) = \omega_1$, specifically $\omega^{\omega_1} = \omega_1$ and $\varepsilon_{\omega_1} = \omega_1$. Furthermore show that $\varphi_{\omega_1}(0) = \omega_1$.

Realize that there was nothing special about the choice of ω. We can apply the same process to ω_1 to get ω_2, and continue doing this to construct larger and larger ordinals (in the sense of cardinality).

Definition 3.45. For an ordinal α we define ω_α as

(i) $\omega_0 \coloneqq \omega$,

(ii) if $\alpha = \beta + 1$ is isolated, then ω_α is the Hartogs number of ω_β,

(iii) if α is a limit ordinal, then $\omega_\alpha \coloneqq \sup\{\omega_\delta \mid \delta < \alpha\}$.

[7] The ordinal ω_1 is also commonly denoted as Ω.

Observation 3.46. *The number ω_α is the first ordinal that is larger (in the sense of cardinality) than all previous ω-numbers.*

This definition foreshadows the section about cardinal numbers, where we will encounter these omega numbers again and explore their properties in depth.

Hartogs numbers also allow us to finally prove that the trichotomy principle implies AC. In fact, this was the original motivation behind Hartogs' theorem.

Theorem 3.47. *The trichotomy principle implies the well-ordering principle.*

Proof. Let x be an arbitrary set, and let η be its Hartogs number. Apply the trichotomy principle to x and η. Exactly one of the following holds:

(a) $x \preceq \eta$, there is an injection $x \to \eta$, or

(b) $\eta \preceq x$, there is an injection $\eta \to x$.

The second case is impossible due to the defining property of η. Hence, there exists an injection $f : x \to \eta$. We can now well-order x by inheriting the order of η by f. □

3.4 Peano Arithmetic

To understand this section, the reader should be familiar with the basic notions of logic, including concepts such as language, theory, model, etc. Explanations of these concepts can be found in the lecture notes [3] for the course NAIL062.

3.4.1 Peano Axioms

Peano Arithmetic, denoted PA, is the standard axiomatic theory of the natural numbers. In ZFC, we have encountered the set of natural numbers, ω, constructed as the set of finite von Neumann ordinals. This is no coincidence; the set ω, together with the restrictions of operations of ordinal arithmetic to ω, serves as the *standard model* for PA, denoted by $\mathcal{N}$.

Our study of ordinal arithmetic in Section 3.2 has already established that these operations, when restricted to finite ordinals, are commutative and satisfy all the familiar properties of elementary arithmetic. The axioms of PA can therefore be seen as a precise, first-order logic attempt to capture the properties of this standard model.

Definition 3.48 (PA, [3])**.** The language of PA is $\mathcal{L}_{PA} = \langle 0, S, +, \cdot, \leq \rangle$ with equality. The base axioms of PA are the following formulas:

$$\begin{aligned} &\neg Sx = 0 && x \cdot 0 = 0 \\ &Sx = Sy \implies x = y && x \cdot Sy = x \cdot y + x \\ &x + 0 = x && \neg x = 0 \implies (\exists y)(x = Sy) \\ &x + Sy = S(x + y) && x \leq y \iff (\exists z)(z + x = y) \end{aligned}$$

These axioms alone yield the much weaker *Robinson Arithmetic* (Q). It cannot prove, for example, the commutativity or associativity of addition or multiplication, or the transitivity of order. To obtain PA, we need to add the *Axiom*

Schema of Induction. That is, for each $\mathcal{L}_{PA}$-formula $\varphi(x, \vec{y})$, the following axiom is added:

$$\big(\varphi(0,\vec{y}) \wedge (\forall x)(\varphi(x,\vec{y}) \to \varphi(Sx,\vec{y}))\big) \implies (\forall x)\varphi(x,\vec{y}) \tag{3.3}$$

Remark. The last axiom schema should seem similar to the induction principle on ω from set theory:

$$(\forall X \subseteq \omega)\Big(\big(0 \in X \wedge (\forall x)(x \in X \Rightarrow x \cup \{x\} \in X)\big) \implies X = \omega\Big).$$

However, the axiom schema of induction is a weaker version, as it is a first-order logic attempt to simulate a second-order logic axiom with an axiom schema. The familiar induction principle could be expressed with the following second-order $\mathcal{L}_{PA}$-formula

$$(\forall X)\Big(\big(X(0) \wedge (\forall x)(X(x) \Rightarrow X(Sx))\big) \implies X = (\forall x)X(x)\Big).$$

By adding it to PA, we would obtain the much stronger second-order theory PA_2.

Here X represents (any) unary relation; that is, a subset of the universe. The important distinction is that (3.3) provides an infinite collection of axioms, one for each subset of the universe that is *definable* by a $\mathcal{L}_{PA}$-formula φ.

This restriction is the source of PA's most profound properties and limitations. For example, PA_2 is categorical; that is, it has only one model (up to isomorphism) — the standard model $\mathcal{N}$. On the other hand, PA allows the existence of other non-standard models.

3.4.2 Models of Arithmetic

We have already mentioned that the *standard model* of PA is the $\mathcal{L}_{PA}$-structure $\mathcal{N} = (\omega, 0^{\mathcal{N}}, S^{\mathcal{N}}, +^{\mathcal{N}}, \cdot^{\mathcal{N}}, \leq^{\mathcal{N}})$, where the domain is the set ω, the interpretation of the symbol '0' is $0^{\mathcal{N}} = \varnothing$, the successor of x is $S^{\mathcal{N}}(x) = x \cup \{x\}$, and $+^{\mathcal{N}}$, $\cdot^{\mathcal{N}}$ and $\leq^{\mathcal{N}}$ are the operations of ordinal arithmetic restricted to ω.

Theorem 3.49. *There exist countable models of* PA *that are not isomorphic to* $\mathcal{N}$.

Proof sketch. By the Compactness Theorem. We extend $\mathcal{L}_{PA}$ with a new constant symbol c. Consider the theory $T = \mathsf{PA} \cup \{c > \bar{n} \,|\, n \in \omega\}$, where $\bar{n}$ is the $\mathcal{L}_{PA}$-term $S(S(\dots S(0)\dots))$ (n times). Any finite subset $T_0 \subset T$ is satisfiable: we take $\mathcal{N}$ as the model and interpret c as a standard natural number larger than any numeral $\bar{n}$ explicitly mentioned in T_0. By the Compactness Theorem, T has a model $\mathcal{M}$. This $\mathcal{M}$ must be a model of PA, but the interpretation of c is a "non-standard" number, an element larger than all standard elements $S^n(0)$. Thus, $\mathcal{M} \ncong \mathcal{N}$. □

All countable non-standard models $\mathcal{M}$ share a common structure: they begin with an initial segment isomorphic to ω (the standard part), which is then followed by a collection of "blocks" of non-standard numbers. This "pathology" of PA is not merely set-theoretic, but also computational.

Theorem 3.50 (Tennenbaum, 1959)**.** *No countable non-standard model of* PA *is recursive.*

This implies that in any non-standard model $\mathcal{M}$, the operations $\oplus$ and $\otimes$ (the interpretations of $+$ and $\cdot$) are not computable functions. Even if the domain of $\mathcal{M}$ is ω, the operations themselves cannot be implemented by an algorithm. The induction schema, while syntactically "weaker" than its second-order counterpart, thus imposes enormous computational complexity on any "non-standard" structure that satisfies it, effectively isolating the standard model as the only computationally tractable one.

3.4.3 Gödel's Incompleteness Theorems

When working with a formal theory, it is natural to ask what statements we can prove from its axioms. If a theory T can prove the sentence ψ, we write $T \vdash \psi$. A theory is *consistent* if it is free from contradictions, meaning it is impossible to prove both a statement ψ and its negation $\neg\psi$ from its axioms; or equivalently, if it has a model. A consistent theory is *complete* if it has an "opinion" on every statement, meaning for every sentence ψ in its language, the theory can prove either ψ or $\neg\psi$. If it cannot do either, it is said to be *incomplete*, and ψ is said to be *independent* in T. Equivalently, ψ is independent in T, if it holds in some models of the theory but does not hold in others.

Probably the most influential result linking these concepts together with PA are the famous Incompleteness Theorems, published by Kurt Gödel [8] [12] in 1931. Veritasium has an amazing video [24] that provides an intuitive explanation of this topic. We provide only a simplified explanation of these profound results; for more details and proofs, refer to [3].

Despite its limitations, PA is a remarkably powerful theory. Its expressive power is sufficient to define all computable (recursive) functions. This strength is the key to PA's own undoing. It allows for the *arithmetization of syntax* (Gödel numbering), whereby the syntax of $\mathcal{L}_{PA}$ (terms, formulas, proofs) can be uniquely encoded as natural numbers. Syntactic operations (like substitution) and relations (like "is a proof of") become recursive functions and relations on these numbers. Crucially, this allows for the creation of a provability predicate.

Definition 3.51. There exists an $\mathcal{L}_{PA}$-formula $\mathrm{Prov}_{\mathsf{PA}}(x)$ such that for any sentence ϕ it holds that $(\mathsf{PA} \vdash \phi) \iff \mathrm{Prov}_{\mathsf{PA}}(\ulcorner\phi\urcorner)$. Here, $\ulcorner\phi\urcorner$ denotes the Gödel number of ϕ. The formula $\mathrm{Prov}_{\mathsf{PA}}(\ulcorner\phi\urcorner)$ is: "there exists x such that x is the Gödel number of a proof of the sentence with Gödel number $\ulcorner\phi\urcorner$."

This predicate allows the theory to "talk about" its own provability, leading directly to sentences that self-reference and assert their own unprovability.

Theorem 3.52 (Gödel's First Incompleteness Theorem, 1931)**.** *If* PA *is consistent, then it is incomplete.*

[8] The life of Kurt Gödel (1906–1978) is a fascinating story. Born in Brno, he left for Vienna at the age of eighteen to study mathematics and logic. At twenty-four, he proved his incompleteness theorems, which formed the basis of his doctoral dissertation. He later emigrated to the United States following the rise of Nazism. Albert Einstein regarded Gödel as the greatest logician since Aristotle and once remarked that the only reason he went to his office was to have the privilege of walking home with Gödel. Yet Gödel's life was not without darkness: he struggled with psychological illness throughout adulthood and ultimately died of self-starvation, driven by the paranoid belief that someone was trying to poison him. Perhaps the most detailed account of Gödel's life (as of the writing of this text) can be found in [2].

Proof sketch. Consider a sentence $\mathfrak{g}$ (the *Gödel sentence*) saying: "there is no x such that x is the Gödel number of a proof of the sentence with Gödel number $\ulcorner \mathfrak{g} \urcorner$." Notice that $(\mathsf{PA} \vdash \mathfrak{g}) \iff \neg \mathrm{Prov}_{\mathsf{PA}}(\ulcorner \mathfrak{g} \urcorner)$. Hence if $\mathsf{PA} \vdash \mathfrak{g}$, then PA is inconsistent. Therefore, if PA is consistent, then $\mathsf{PA} \nvdash \mathfrak{g}$, and it is incomplete. □

As a corollary of this theorem, Gödel achieved his second result.

Theorem 3.53 (Gödel's Second Incompleteness Theorem, 1931)**.** PA *cannot prove its own consistency.*

Proof sketch. Let $\mathrm{Con}(\mathsf{PA})$ be the $\mathcal{L}_{PA}$-sentence $\neg \mathrm{Prov}_{\mathsf{PA}}(\ulcorner \bot \urcorner)$ (where $\bot$ is a contradiction, e.g., $0 = S0$). That is, $\mathrm{Con}(\mathsf{PA})$ is true if and only if PA is consistent. Because Gödel formalized the entire proof of the previous theorem in PA (using Gödel numbers), his first theorem can be expressed as

$$\mathsf{PA} \vdash \big(\mathrm{Con}(\mathsf{PA}) \implies \neg \mathrm{Prov}_{\mathsf{PA}}(\ulcorner \mathfrak{g} \urcorner)\big).$$

This together with the equivalence $(\mathsf{PA} \vdash \mathfrak{g}) \iff \neg \mathrm{Prov}_{\mathsf{PA}}(\ulcorner \mathfrak{g} \urcorner)$ gives

$$\mathsf{PA} \vdash \big(\mathrm{Con}(\mathsf{PA}) \implies \mathfrak{g}\big).$$

Now, suppose for contradiction that PA could prove its own consistency. Combining this with the last formula gives $\mathsf{PA} \vdash \mathfrak{g}$, but this is a contradiction, since (the end of the previous proof) if PA is consistent, then $\mathsf{PA} \nvdash \mathfrak{g}$. □

Gödel's original formulation of these theorems did not, in fact, talk about PA, but about a system he called P, a close relative of PA. Gödel then had to make a philosophical assumption. He argued that any other system "related" to, and at least as strong as P (and therefore capable of arithmetic), would also be capable of producing a Gödel sentence $\mathfrak{g}$; thus, his incompleteness theorems would apply to this system as well. This was a strong, intuitive argument, but he could not formally prove it.

The missing piece was provided in 1936 by Alan Turing [23], who formalized the notion of computability using the Turing machine, which made a formal proof of Gödel's conjecture possible.

Theorem 3.54 (Generalized Gödel's Incompleteness Theorems)**.** *For any consistent, recursively axiomatized theory T, it holds that:*

(1) *If T is an extension of Robinson arithmetic Q, then T is incomplete.*

(2) *If T is an extension of Peano arithmetic PA, then T cannot prove its own consistency.*

Remark. Recursively axiomatized means that there is an algorithm (Turing machine) that, for every input formula φ, halts and answers whether φ is an axiom of T. The condition that T is an extension of Q (or PA) essentially means that T is at least as powerful as Q (or PA). For example, PA is an extension of Q.

Corollary 3.55. *It is impossible to prove the consistency of* ZFC *inside* ZFC.

An example of an independent statement in ZFC is the famous continuum hypothesis CH, claiming that there is no set x such that $\omega \prec x \prec \mathcal{P}(\omega)$. Similarly, AC can be shown to be independent in ZF, meaning that if ZF is consistent, then ZFC is as well.

In 1940, Gödel showed that neither AC can be disproved from ZF, nor CH from ZFC, by constructing the *constructible universe*, a model of ZF in which both AC and CH hold. This model begins with the empty set and adds only those sets that are definable from previous ones, thus forming the minimal universe compatible with the axioms. Later, in 1963, Paul Cohen showed that CH cannot be proved from ZFC by developing the method of *forcing*, which allowed him to construct a model of ZFC in which CH fails. Through a different forcing argument, he likewise obtained a model of ZF that violates AC.

3.4.4 Consistency and the Connection with ε_0

Gödel's second theorem seems to place us in a difficult position: a consistency proof for PA must employ principles that transcend PA itself. While ZFC is far stronger than PA and easily proves Con(PA) (by exhibiting the model $\mathcal{N}$), this isn't a very "unilluminating" result. PA is a "finitary" theory, while ZFC is a wildly "infinitary" theory (it assumes the existence of various vast infinities). By proving the consistency of PA in ZFC, we base our proof on the assumption that ZFC is consistent. It would be better to find a weaker system that is still capable of proving Con(PA).

Theorem 3.56 (Gentzen's Consistency Proof,[9] 1936)**.** *The consistency of* PA *is provable in Primitive Recursive Arithmetic* PRA *(which by itself is weaker than* PA*), augmented with a schema for transfinite induction up to the ordinal* ε_0*.*

Gentzen's proof precisely identified the principle transcending PA that is required to prove its consistency. Recall from Section 3.3.1 that ε_0 is the first fixed point of the ordinal function $\alpha \mapsto \omega^\alpha$, the limit of the sequence $\omega, \omega^\omega, \omega^{\omega^\omega}, \ldots$ Gentzen's result, $\mathsf{PRA} + \mathrm{TI}(\varepsilon_0) \vdash \mathrm{Con}(\mathsf{PA})$, thus establishes two facts:

(a) Proving the consistency of PA does not require the full power of ZFC; only transfinite induction up to a countable ordinal. That is, the assumption that ε_0 contains no infinite decreasing chains.

(b) The principle of transfinite induction up to ε_0, $\mathrm{TI}(\varepsilon_0)$, must be unprovable in PA (lest PA prove its own consistency).

Gentzen also showed that using any smaller ordinal $\alpha < \varepsilon_0$ is not enough. This calibrates the strength of PA with extraordinary precision. The collected strength of PA's infinite induction schema is exactly equivalent to the single principle of transfinite induction up to (but not including) ε_0. This is formalized in the concept of the *proof-theoretic ordinal*.

Theorem 3.57. *The* proof-theoretic ordinal *of* PA *is* $|\mathsf{PA}| = \varepsilon_0$.

This theorem has a twofold meaning that we can understand intuitively:

[9] See [20] for a modern version of the proof. Moreover, [7] provides an alternative view on this result, and talks about the consistency of PA in general.

(a) What PA *can* prove: PA is strong enough to prove the well-foundedness of any recursive well-ordering $<_R$ on ω with order-type $\alpha < \varepsilon_0$.

(b) What PA *cannot* prove: PA is *not* strong enough to prove the well-foundedness of any recursive well-ordering $<_R$ on ω with order-type $\alpha \geq \varepsilon_0$.

Here, "recursive" means that there exists an algorithm that can answer whether $x <_R y$ or $y <_R x$ for all $x, y \in \omega$. Well-foundedness is the arithmetical statement that every nonempty subset of ω has a minimal element. Proving well-foundedness is thus equivalent to proving that transfinite induction "works" for that ordering (as there cannot be any infinite decreasing chains).

A point of confusion here might be the fact that any well-ordering is well-founded. But PA does not know that $<_R$ is a well-ordering; it only receives an "object," $<_R$, together with the instructions: "prove that the ordering you received is well-founded."

Therefore, PA can formalize proofs by transfinite induction up to any ordinal $\alpha < \varepsilon_0$, but it cannot justify the principle of transfinite induction up to ε_0 itself.

3.4.5 Limits of Predicative Mathematics

The discovery of set-theoretic paradoxes (such as Russell's and Burali-Forti's) in the early 20th century triggered the so-called *Grundlagenkrise*, or foundational crisis, in mathematics. The naive assumption that any property $\phi(x)$ could define a set $\{x \mid \phi(x)\}$ was shown to lead to contradictions. This prompted a range of philosophical responses, but most mathematicians eventually turned to the axiomatic framework of ZFC, which resolved the paradoxes by carefully restricting what counts as a set.

However, a significant objection came from mathematicians like Poincaré and, most notably, Hermann Weyl. They argued that the core problem was the use of *impredicative definitions*—definitions that define an object S by quantifying over a totality T that already includes S. This went directly against ZFC, as it uses impredicative definitions all the time. For example, the supremum of a set is defined as its least upper bound. In this case, the totality is the set of all upper bounds, and since the supremum is itself an upper bound, it is a member of that totality.

Weyl argued that such definitions were circular and potentially dangerous. He decided to rebuild mathematical analysis on a "safe" *predicative* basis, starting with his 1918 paper "The Continuum." He succeeded in developing a significant portion of classical analysis, but was unable to replicate everything. The big question became: "How much of mathematics can we *actually* recover using only predicative methods?" In the 1960s, Solomon Feferman and Kurt Schütte independently found the precise answer. We provide an intuitive interpretation of their result.

Theorem 3.58 (Feferman–Schütte, c. 1965)**.** *The proof-theoretic ordinal of "predicative mathematics" is the* Feferman–Schütte ordinal*, denoted* Γ_0*. It is the first fixed point of the function* $\alpha \mapsto \varphi_\alpha(0)$ *(see Section 3.3.2).*

This shows that Γ_0 is the first ordinal that cannot be proven to be well-founded by predicative means, just as ε_0 is the first ordinal that cannot be proven well-founded by the "finitistic" means of PA.

Remark. It should be noted that not all mathematicians agree on what "predicative mathematics" means exactly. It would be more accurate to say that Feferman and Schütte showed that Γ_0 is the first ordinal that cannot be proven well-founded by *certain* predicative means, and most people agree that those means are a reasonable interpretation of predicativity. This was later formalized in the 1970s by Friedman and Simpson in a formal system called ATR_0 (Arithmetical Transfinite Recursion). So what the theorem above really says is that the proof theoretic ordinal of ATR_0 is Γ_0.

The discipline of finding the proof-theoretic ordinals of theories is called *Ordinal Analysis*. For an introduction to the topic, I recommend [20]. Other notable sources are [8], [21] and [18].

3.5 Applications of Countable Ordinals

Gödel's independent sentences, $\mathfrak{g}$ and Con(PA), are meta-mathematical statements, not "natural" theorems of number theory. For decades, it was an open question whether any "ordinary" theorem of arithmetic or combinatorics was unprovable in PA, leading to speculation that Gödel's Incompleteness Theorem would not have meaningful implications to practical mathematics.

However, in 1977, Parris and Harrington [19] showed that a very natural variation of Ramsey's Theorem was true but not provable in PA. Five years later, in 1982, Kirby and Paris [17] showed that Goodstein's theorem, a statement purely about sequences of natural numbers, cannot be proven in PA either.

The second theorem presented in Kirby and Paris's 1982 paper establishes an analogous result, this time showing that a statement about the Hydra game is unprovable in PA.

3.5.1 Goodstein Sequences

Definition 3.59. The hereditary base-n form of a natural number m is achieved by first writing m in base n, and then applying the procedure inductively to each exponent until there are no numbers larger than n.

Example. $100_2 = 64 + 32 + 4 = 2^6 + 2^5 + 2^2 = 2^{2^2+2} + 2^{2^2+1} + 2^2$.

The Goodstein sequence of a natural number $m > 0$ is generated as follows.

- Start with $m_0 = m$.
- To get m_1, write m_0 in hereditary base-2 notation, replace all 2s with 3s, and then subtract 1 from the result.
- To get m_{n+1} from m_n, write m_n in hereditary base-$(n+2)$ notation, replace each occurrence of $(n+2)$ with $(n+3)$, and subtract 1.
- If ever $m_n = 0$, then $m_{n+1} = 0$.

For example, when we start with $m_0 = 3$, we get the sequence:

$$\begin{aligned} m_0 &= 2^1 + 1 && = 3 \\ m_1 &= 3^1 + 1 - 1 = 3^1 && = 3 \end{aligned}$$

$$
\begin{aligned}
m_2 &= 4^1 - 1 = 3 && = 3 \\
m_3 &= 3 - 1 = 2 && = 2 \\
m_4 &= 2 - 1 = 1 && = 1 \\
m_5 &= 1 - 1 = 0 && = 0
\end{aligned}
$$

Notice that at step $m_2 \to m_3$, the base used (4 at m_2) exceeded the value of $m_2 = 3$, which caused the sequence to start decreasing, and it eventually terminated. Does it always terminate? Let's try again, this time with $m_0 = 29$.

$$
\begin{aligned}
m_0 &= 2^{2^2} + 2^{2+1} + 2^2 + 1 && = 29 \\
m_1 &= 3^{3^3} + 3^{3+1} + 3^3 && \sim 8 \cdot 10^{12} \\
m_2 &= 4^{4^4} + 4^{4+1} + 4^4 - 1 = && \sim 10^{154} \\
&= 4^{4^4} + 4^{4+1} + 3 \cdot 4^3 + 3 \cdot 4^2 + 3 \cdot 4 + 3 && \\
m_3 &= 5^{5^5} + 5^{5+1} + 3 \cdot 5^3 + 3 \cdot 5^2 + 3 \cdot 5 + 2 && \sim 10^{2200} \\
m_4 &= 6^{6^6} + 6^{6+1} + 3 \cdot 6^3 + 3 \cdot 6^2 + 3 \cdot 6 + 1 && \sim 10^{36305}
\end{aligned}
$$

This does not look like it will terminate. Let's try it one more time, this time with $m_0 = 4$, to really understand what is going on.

$$
\begin{aligned}
m_0 &= 2^2 && = 4 \\
m_1 &= 3^3 - 1 = 2 \cdot 3^2 + 2 \cdot 3 + 2 && = 26 \\
m_2 &= 2 \cdot 4^2 + 2 \cdot 4 + 1 && = 41 \\
m_3 &= 2 \cdot 5^2 + 2 \cdot 5 && = 60 \\
&\vdots && \vdots \\
m_9 &= 2 \cdot 11^2 + 11 && = 253 \\
m_{10} &= 2 \cdot 12^2 + 12 - 1 = 2 \cdot 12^2 + 11 && = 299 \\
&\vdots && \vdots \\
m_{22} &= 2 \cdot 24^2 - 1 = 24^2 + 23 \cdot 24 + 23 && = 1151 \\
&\vdots && \vdots \\
m_{B-2} &= 2 \cdot B^1 && = 2B > 10^{10^8} \\
m_{B-1} &= 2 \cdot (B+1)^1 - 1 = (B+1)^1 + B && = 2B + 1 \\
m_B &= (B+2)^1 + B - 1 && = 2B + 1 \\
&\vdots && \vdots \\
m_{B+k} &= (B+k+2)^1 + B - (k+1) && = 2B + 1 \\
&\vdots && \vdots \\
m_{2B-2} &= (2B)^1 + 1 && = 2B + 1 \\
m_{2B-1} &= (2B+1)^1 && = 2B + 1 \\
m_{2B} &= (2B+2)^1 - 1 = 2B + 1 && = 2B + 1 \\
m_{2B+1} &= 2B + 1 - 1 && = 2B \\
&\vdots && \vdots
\end{aligned}
$$

$$
\begin{aligned}
m_{2B+k} &= 2B+1-k && = 2B-(k-1)\\
&\vdots && \vdots\\
m_{3B} &= 2B+1-B=1 && = B+1\\
&\vdots && \vdots\\
m_{4B} &= 2B+1-2B=1 && = 1\\
m_{4B+1} &= 0 && = 0
\end{aligned}
$$

This demonstrates that when we start at $m_0 = 4$, the sequence first rises to index $B-2$, where it attains the value $2 \cdot B^1$ (current base is B). In the next step, this is decomposed to $(B+1)^1 + B$ (this is also the maximum value). The sequence lingers here for the next B steps (each base change raises the value by 1, and this is immediately subtracted), and afterward, it begins its long descent. Finally, at index $4B+1$, it reaches zero. The exact value of B is $3 \cdot 2^{402\,653\,209} - 1$, and the length of this sequence is $4B + 2 = 3 \cdot 2^{402\,653\,211} - 2$.

Goodstein's theorem claims that a similar fate awaits every Goodstein sequence, no matter how large the starting value might be.

Theorem 3.60 (Goodstein, 1944)**.** *For every natural number m, there exists a natural number n such that $m_n = 0$.*

Definition 3.61 (Change of base function)**.** For a natural number $b \geq 2$ we define the functions

(a) *change of base function* $R_b : \omega \to \omega$ as the function that takes a natural number n and replaces each b in the hereditary base-b representation of n with $b+1$.

(b) *ω-change of base function* $R_b^\omega : \omega \to \varepsilon_0$ as the function that takes a natural number n and replaces each b in the hereditary base-b representation of n with ω.

Formally, we define $R_b(0) = R_b^\omega(0) = 0$ and for $n > 0$ expressed in base-b as

$$n = \sum_{i=0}^{k} b^i \cdot p_i$$

we let

$$R_b(n) = \sum_{i=0}^{k} b^{R_b(i)} \cdot p_i, \qquad R_b^\omega(n) = \sum_{i=0}^{k} b^{R_b^\omega(i)} \cdot p_i.$$

Example. For example:

- $R_2(29) = R_2(2^{2^2} + 2^{2+1} + 2^2 + 1) = 3^{3^3} + 3^{3+1} + 3^3 + 1$,
- $R_2^\omega(29) = R_2^\omega(2^{2^2} + 2^{2+1} + 2^2 + 1) = \omega^{\omega^\omega} + \omega^{\omega+1} + \omega^\omega + 1$

Remark. Since each $n \in \omega$ is finite, the hereditary base-b representation of n is finite, and thus $R_b^\omega(n)$ also contains only finitely many occurrences of ω. Therefore $R_b^\omega(n) < \varepsilon_0$ for each $b \geq 2$ and $n \in \omega$.

Observation 3.62. *The terms of a Goodstein sequence starting in m could be defined as $m_{n+1} = R_{n+2}(m_n) - 1$.*

Lemma 3.63. *For every $b \geq 2$ and $n \geq 0$ it holds that $R^\omega_b(n+1) > R^\omega_b(n)$.*

Intuition. This should seem obvious, for example when $b = 2$ and $n = 13$ we get

$$R^\omega_2(13) = R^\omega_2(2^{2+1} + 2^2 + 1) = \omega^{\omega+1} + \omega^\omega + 1$$
$$R^\omega_2(13+1) = R^\omega_2(2^{2+1} + 2^2 + 2) = \omega^{\omega+1} + \omega^\omega + \omega,$$

Proof. Let $b \geq 2$ be given; we prove the claim using induction on n. If $n = 0$, then we have $R^\omega_b(1) = 1 > 0 = R^\omega_b(0)$. If $n > 0$, then let the following be the base-b expansions of n and $n+1$:

$$n = b^{c_0} \cdot p_0 + b^{c_1} \cdot p_1 + \cdots + b^{c_k} \cdot p_k$$
$$n+1 = b^{d_0} \cdot q_0 + b^{d_1} \cdot q_1 + \cdots + b^{d_l} \cdot q_l$$

where all p_i and q_j are nonzero, and $c_0 > c_1 > \cdots > c_k$ and $d_0 > d_1 > \cdots > d_l$. Theorem 3.32, or more precisely, part two of Theorem 3.30, describes the two possible conditions that these expansions must satisfy in order for one of them $(n+1)$ to be larger than the other (n). Notice that when we apply R^ω_b to these expansions, then (assuming the claim already holds for all $k < n$), the conditions remain unchanged; therefore $R^\omega_b(n+1) > R^\omega_b(n)$. □

We can now prove Goodstein's theorem.

Proof of Theorem 3.60 (cf. [22]). Let $m = m_0$ be given. We will define a sequence of ordinals $\mu_n < \varepsilon_0$ satisfying

$$\mu_n > 0 \implies \mu_{n+1} < \mu_n \qquad \text{and} \qquad \mu_n > 0 \iff m_n > 0.$$

Since ε_0 is well-founded, it does not admit infinite strictly decreasing sequences; thus, there exists an index k at which $\mu_k = 0$, and therefore also $m_k = 0$.

Define $\mu_n \coloneqq R^\omega_{n+2}(m_n)$. Clearly $\mu_n > 0 \iff m_n > 0$, so it remains to show that μ_n is decreasing. If $\mu_n > 0$, then

$$\begin{aligned}
\mu_{n+1} &= R^\omega_{n+3}(m_{n+1}) \\
&= R^\omega_{n+3}(R_{n+2}(m_n) - 1) \\
&< R^\omega_{n+3}(R_{n+2}(m_n)) \\
&= R^\omega_{n+2}(m_n) = \mu_n.
\end{aligned}$$

The inequality holds by Lemma 3.63 since if $\mu_n > 0$, then $m_n > 0$ and so $R_{n+2}(m_n) \geq m_n > 0$. The equality after the inequality is a trivial property of the base change functions ($n+2 \to n+3 \to \omega$ is the same as $n+2 \to \omega$). □

Extended Goodstein's theorem In the version of Goodstein sequences presented above, we started with base $b_0 = 2$ for m_0, then changed it to $b_1 = 3$ for m_1, and in general worked with base $b_n = n + 2$ for m_n. This can be generalized by considering any non-decreasing sequence $2 \leq b_0 \leq b_1 \leq \cdots$ of bases and defining the term m_{n+1} from m_n by expressing m_n in hereditary base-b_n notation,

replacing each occurrence of b_n with b_{n+1}, and subtracting one. It is not hard to modify the proof above to show that this sequence still terminates.

This extended version is, in fact, the one Goodstein originally considered in [13], and he proved that it is equivalent to the claim that ε_0 is well-founded. We have mentioned in Section 3.4.4 that PA cannot justify the well-foundedness of ε_0; so does that mean that Goodstein showed that the extended Goodstein's theorem cannot be proved in PA? Unfortunately, he didn't, because the extended theorem is not formalizable in PA as it cannot represent arbitrary infinite sequences.

The simple version we considered can however be formalized in PA easily (as we are working with only one specific sequence), and it seems much "tamer" than the extended version. For a long time it was unknown whether the simple version could be proved without using tools beyond the reach of PA (such as the well-foundedness of ε_0). Almost forty years later, Kirby and Paris [17] showed that no such finitary proof is possible.

Theorem 3.64 (Kirby–Paris, 1982)**.** *Goodstein's theorem is true, but unprovable in Peano arithmetic.*

Intuition. Even though for every fixed natural number m we have

$$\mathsf{PA} \vdash (\exists n)(\overline{m}_n = 0),$$

where $\overline{m} = S(S(\dots S(0)\dots))$ repeated m times, (for each fixed m, PA can verify the finite descent of the Goodstein sequence by explicit computation, showing that it terminates after some finite number n of steps), it also holds that

$$\mathsf{PA} \nvdash (\forall m)(\exists n)(m_n = 0).$$

Proof sketch. Kirby and Paris started with a statement of the form

$$(\forall a)(\forall b)(\exists c)\, \varphi(a,b,c)$$

that was known to be independent[10] in PA. Then they constructed a nonstandard model M of PA containing a nonstandard $b_0 \in M$ such that

$$M \models \neg(\exists y)\, \varphi(1, b_0, y).$$

That is, the above-mentioned independent statement does not hold in this model. To finish the proof, they showed that if Goodstein's theorem could be proved in PA, then there would exist a (very large) number $e \in M$ (satisfying $m_e = 0$ for some carefully chosen m defined using the nonstandard b_0) such that $\varphi(1, b_0, e)$ holds (inside M), which is a contradiction. □

This result should be surprising. As shown in [22], PA is equivalent to the theory of finite sets; that is, ZFC with the axiom of infinity replaced by the axiom "there are no limit ordinals." From this, one can prove that all sets are finite. The Kirby–Paris theorem asserts that accepting an axiom about infinite sets changes what we can prove about finite sets.

[10] The formula $\varphi(a,b,c)$ talks about certain large ordinals $\alpha < \varepsilon_0$, utilizing the fact that these numbers have a finite hereditary Cantor normal form, and can therefore be formalized in PA.

3.5.2 The Hydra Game

There are different versions of the Hydra game; the one we will focus on was presented by Kirby and Paris in [17], (1982).

A *hydra* is a finite rooted tree, usually drawn with the root at the bottom. A *head* of the hydra is a leaf together with its attached edge. A *battle* between Hercules and a given hydra is divided into stages, starting at stage one. During stage n, Hercules chops off one head of the hydra. The hydra then grows n new "heads" in the following manner:

- From the node that used to be attached to the head which was just chopped off, move along the length of one edge towards the root; that is, move to the grandparent of the chopped off node.
- From this node, sprout n replicas of that part of the hydra (after decapitation), which is "above" the edge just traveled.
- If the head just chopped off was attached to the root, no new head is grown.

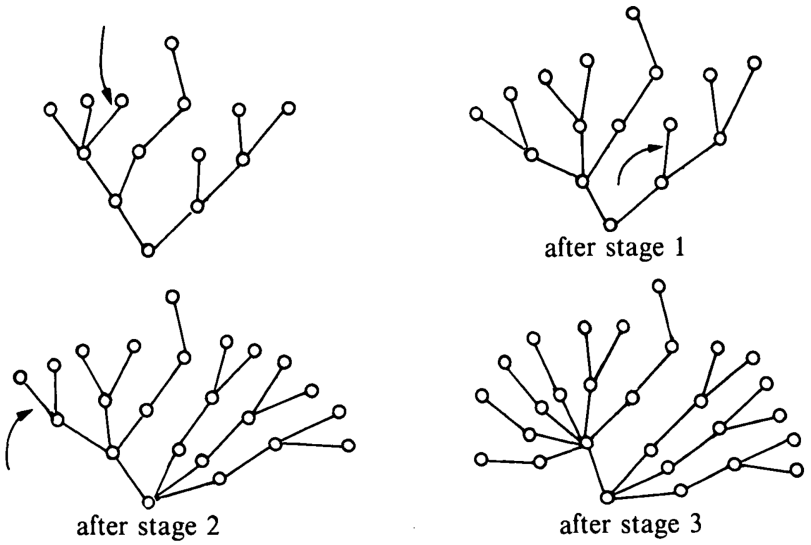

Figure 4: The Hydra game, as presented by Kirby and Paris [17], if at each stage Hercules decides to chop off the head marked with an arrow.

Hercules *wins* if after some finite number of stages, the hydra is dead; that is, nothing is left of the hydra but its root. A *strategy* is a function which determines for Hercules which head to chop off at each stage of any battle. A strategy is *winning* if it ensures that Hercules wins against every hydra. It is not hard to show that a winning strategy exists (for example, always targeting one of the highest positioned heads). More surprisingly, Hercules cannot help winning:

Theorem 3.65. *Every strategy is a winning strategy.*

Proof. The idea of the proof is very similar to that of Goodstein's theorem. We create a sequence of ordinal numbers $\mu_n < \varepsilon_0$ that is strictly decreasing, and its elements are positive if and only if the hydra is still alive. Since ε_0 is well-founded,

it admits no infinite decreasing chains; thus, the sequence must eventually terminate, and the hydra will die with it.

We can assign to each node of the hydra an ordinal $\alpha < \varepsilon_0$ as follows:

- assign 0 to each leaf,
- and to every other node assign $\omega^{\alpha_1} + \cdots + \omega^{\alpha_k}$ where $\alpha_1 \geq \cdots \geq \alpha_k$ are the ordinals assigned to the nodes immediately "above." ($\omega^0 = 1$).

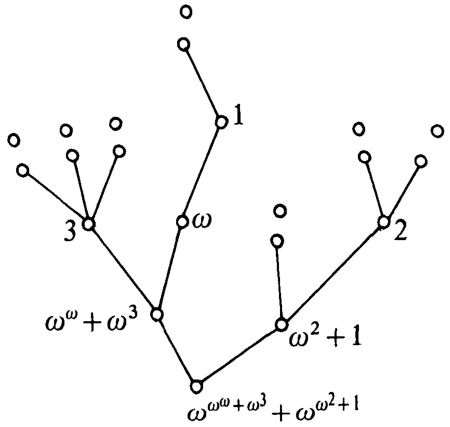

Figure 5: Ordinal assignments of the hydra from the example above; adapted from Kirby and Paris [17].

The ordinal of a hydra is the ordinal assigned to its root. Notice that a hydra is dead if and only if its ordinal is zero. For any strategy σ, we can define a function $H_\sigma(\alpha, n) : \varepsilon_0 \times \omega \to \varepsilon_0$ that maps the ordinal of the hydra represented by α, together with a stage number n, to the ordinal of the hydra that results from Hercules chopping off the head of α as specified by the strategy σ at stage $n - 1$ (that is, n new "heads" are grown).

To prove the theorem, one only has to show that for any strategy σ, any $0 < \alpha < \varepsilon_0$, and any $n \in \omega$, it holds that $H_\sigma(\alpha, n) < \alpha$. □

Exercise 13. Finish the proof by showing that $H_\sigma(\alpha, n) < \alpha$.

Kirby and Paris then proved a second independence statement.

Theorem 3.66 (Kirby–Paris, 1982)**.** *The statement "every recursive strategy is a winning strategy" is not provable in Peano arithmetic.*

Remark. "Recursive strategy" means any strategy that can be implemented by an algorithm. The reason why the theorem does not simply say "every strategy" is that it is not possible to represent an arbitrary infinite object in PA. The restriction to recursive strategies ensures that each one of them can be formalized by a (finite) Turing machine.

3.5.3 Fundamental Sequences

Definition 3.67. The *fundamental sequence* of a countable limit ordinal α is an increasing sequence $\alpha[0] < \alpha[1] < \ldots$ such that $\alpha = \sup\{\alpha[n] \mid n < \omega\}$.

Remark. Note that Lemma 3.35 implies that no uncountable ordinal can have a fundamental sequence.

It is clear from the definition that the same ordinal α may have multiple fundamental sequences, but there is usually a "standard" one.

Definition 3.68. The following is a common assignment of fundamental sequences to limit ordinals $\alpha \leq \varepsilon_0$ (the first fixed point of $\xi \mapsto \omega^\xi$). Corollary 3.31 claims that every $\alpha < \varepsilon_0$ can be expressed as

$$\alpha = \omega^{\gamma_0} + \omega^{\gamma_1} + \cdots + \omega^{\gamma_k},$$

where $\gamma_0 \geq \gamma_1 \geq \cdots \geq \gamma_k$ (by expanding the Cantor normal form of α). Since $\alpha < \varepsilon_0$, each exponent γ_i satisfies $\gamma_i < \alpha$. Also, note that since α is limit, we have that $\gamma_k \neq 0$. We define $\alpha[n]$ inductively as

(i) $\omega^{\gamma+1}[n] \coloneqq \omega^\gamma \cdot (n+1)$,

(ii) $\omega^\gamma[n] \coloneqq \omega^{\gamma[n]}$ for limit ordinals γ

(iii) $(\omega^{\gamma_0} + \cdots + \omega^{\gamma_k})[n] \coloneqq \omega^{\gamma_0} + \cdots + (\omega^{\gamma_k}[n])$ for $\gamma_0 \geq \cdots \geq \gamma_k$,

(iv) $\varepsilon_0[n] = \gamma_n$, where $\gamma_0 = 1$ and $\gamma_{n+1} = \omega^{\gamma_n}$. That is, $\gamma_n = {}^n\omega$ is a ω-tower of height n.

Example. Some simple fundamental sequences are

- $\omega[n] = \omega^0 \cdot (n+1) = n+1$,
- $\omega^\omega[n] = \omega^{\omega[n]} = \omega^{n+1}$,
- $\omega^{\omega+5}[n] = \omega^{\omega+4} \cdot (n+1)$,
- $(\omega^{\omega+2} + \omega^\omega \cdot 5)[n] = \omega^{\omega+2} + \omega^\omega \cdot 4 + (\omega^\omega[n]) = \omega^{\omega+2} + \omega^\omega \cdot 4 + \omega^{n+1}$

Fundamental sequences can also be assigned to limit ordinals larger than ε_0, but this becomes much more complicated. The next common approach after Cantor normal form is to utilize the Veblen function described in Section 3.3.2. One can show that all ordinals $\alpha < \Gamma_0$ have a unique normal form in terms of Veblen functions; for details, see Section 8 of [10]. This normal form can then be utilized to define fundamental sequences for all limit ordinals $\alpha < \Gamma_0$. If you are interested in how the formulas look, see [27].

3.5.4 The Fast-Growing Hierarchy

One of the use cases of fundamental sequences is the definition of hierarchies of functions $f_\alpha : \omega \to \omega$, where each function grows faster than the previous one.

Definition 3.69 (Fast-growing hierarchy[11])**.** For ordinals $\alpha \leq \varepsilon_0$, we define functions from natural numbers to natural numbers $f_\alpha : \omega \to \omega$ as follows:

(i) $f_0(n) \coloneqq n+1$,

[11] Numberphile has a VIDEO where they explore the growth rates of some extremely fast-growing functions (Graham's iteration and the TREE sequence).

(ii) $f_{\alpha+1}(n) \coloneqq f_\alpha^n(n) = f_\alpha(f_\alpha(\dots f_\alpha(n)\dots))$, where f_α is composed n times,

(iii) $f_\alpha(n) \coloneqq f_{\alpha[n]}(n)$ for limit ordinals α.

Remark. Martin Löb and Stanley Wainer introduced this hierarchy in 1970s as a generalization of the Grzegorczyk hierarchy, which only considered $\alpha < \omega$.

Remark. The ordinal ε_0 is not important in the definition; we could use any other large countable ordinal μ if we had fundamental sequences for all limit ordinals $\alpha \leq \mu$. Also, note that the values of the functions f_α may differ based on the chosen fundamental sequences for limit ordinals. The idea is that the values will be asymptotically the same.

Example. Some fast-growing hierarchy functions are

- $f_1(n) = 2n, \quad f_2(n) = 2^n \cdot n,$
- $f_3(n) > {}^n2$, where ${}^n2$ is a 2-tower of height n; this is called *tetration*.
- $f_\omega(n) > A(n,n)$, where A is the *Ackermann function*.[12]

Definition 3.70. A function $f : \omega \to \omega$ is

(a) *total* if it is defined for every $n \in \omega$; *partial* otherwise,

(b) *recursive* (or *computable*) if there exists an algorithm that for any given input n halts precisely when $f(n)$ is defined and outputs $f(n)$,

(c) *primitive recursive* if there exists an algorithm that does not use recursion (it only uses loops and conditions) and for any given input n halts precisely when $f(n)$ is defined and outputs $f(n)$.

It is important to realize that all the functions f_α defined above are recursive, as for each of them, there is a straightforward algorithm to compute the value $f_\alpha(n)$ for any n by following α down to smaller ordinals using the fundamental sequences $\alpha[n]$. Even if the ordinal α is infinite, a Turing machine set to compute f_α would eventually find $f_\alpha(n)$ in a finite amount of time.

We have already mentioned that we can define fundamental sequences for ordinals $\alpha \leq \Gamma_0$ (and there are ways to go beyond), and the functions f_α would remain recursive as long as we have a well-defined, recursive method for choosing the sequence $\alpha[n]$. This raises the question: "How far can we go?" Eventually, it must become impossible to choose $\alpha[n]$ in a recursive manner because there are only countably many recursive functions (each corresponds to a Turing machine, which can be represented as a finite sequence of natural numbers, and there are only countably many of those), while there are uncountably many countable ordinals, as we have shown in Section 3.3.3.

Therefore, at some point, we will reach a countable "non-recursive" ordinal Λ for which f_Λ can no longer be recursive. The first such ordinal is called the *Church–Kleene ordinal* and it is denoted by ω_1^{CK}. All ordinals $\alpha < \omega_1^{\mathrm{CK}}$ are *recursive*; that means there exists a recursive well-ordering $<_\alpha$ of ω with type α. How big is the Church–Kleene ordinal? It is certainly much, much larger than Γ_0, but it is still countable, so nothing compared to ω_1.

[12] Computerphile has a VIDEO about the Ackermann function.

The Ackermann Function

The previously mentioned Ackermann function is a famous total recursive function that is not primitive recursive. It is defined as follows.

Algorithm 1: Ackermann function $A(m,n)$

```
Function Ackermann(m,n):
    if m = 0 then
        return n + 1;
    else
        if n = 0 then
            return Ackermann(m − 1,1);
        else
            return Ackermann(m − 1,Ackermann(m,n − 1));
```

Exercise 14. Convince yourself that the algorithm above always halts.

Definition 3.71. We say that a function $f : \omega \to \omega$ dominates a function $g : \omega \to \omega$ if, for all sufficiently large n, we have $f(n) > g(n)$.

The fast-growing hierarchy is sometimes referred to as the Grzegorczyk hierarchy due to the following theorem.

Theorem 3.72 (Grzegorczyk, 1953)**.** *Any primitive recursive function is eventually dominated by some f_n for $n \in \omega$.*

This means that the functions f_n for $n \in \omega$ measure the growth rates of all primitive recursive functions.

Corollary 3.73. *Since the Ackermann function is not primitive recursive, $A(n,n)$ dominates all of these. It can also be shown that $A(n,n)$ is dominated by $f_\omega(n)$, so the Ackermann function lies at the very edge between primitive and non-primitive recursiveness.*

The Hardy Hierarchy

Definition 3.74 (Hardy hierarchy, 1904)**.** For ordinals $\alpha \leq \varepsilon_0$, we define functions from natural numbers to natural numbers $H_\alpha : \omega \to \omega$ as follows:

(i) $H_0(n) \coloneqq n$,

(ii) $H_{\alpha+1}(n) \coloneqq H_\alpha(n+1)$,

(iii) $H_\alpha(n) \coloneqq H_{\alpha[n]}(n)$ for limit ordinals α.

Example. Some simple Hardy functions are

- $H_1(n) = H_0(n+1) = n+1$,
- $H_k(n) = n+k$,
- $H_\omega(n) = H_{\omega[n]}(n) = H_{n+1}(n) = n + (n+1) = 2n+1$.

Exercise 15. Try computing $H_{\omega+\omega}(n), H_{\omega\cdot k}(n)$ and $H_{\omega\cdot\omega}$.

The Hardy hierarchy seems to be much slower than the fast-growing hierarchy; they are in fact related by $f_\alpha \sim H_{\omega^\alpha}$ for all $\alpha \leq \varepsilon_0$. However, notice that the Hardy hierarchy "catches up" at $\alpha = \varepsilon_0$ (since $\varepsilon_0 = \omega^{\varepsilon_0}$) in the sense that

$$f_{\varepsilon_0}(n-1) \leq H_{\varepsilon_0}(n) \leq f_{\varepsilon_0}(n+1).$$

This means that the two hierarchies can often be treated as equal.

Theorem 3.75 (Schwichtenberg–Wainer, c. 1972)**.** *The total recursive functions that can be proved total by Peano arithmetic are exactly those that are eventually dominated by some f_α (or equivalently some H_α) for $\alpha < \varepsilon_0$.*

Remark. By "PA can prove that f is total," we mean that PA can prove that the algorithm defining f always terminates.

Corollary 3.76. PA *cannot prove that neither f_{ε_0} nor H_{ε_0} are total.*

This is a "computational" counterpart of Gentzen's consistency proof we saw in Section 3.4.4. Gentzen showed that PA cannot prove that ε_0 is well-founded. Intuitively, it should not be able to prove that the recursive definition of H_{ε_0} is total (that every recursive call eventually reaches the base case H_0).

Connection to Goodstein's Theorem

Definition 3.77 (Goodstein function)**.** The function $\mathcal{G} : \omega \to \omega$ mapping each natural number m to the length of the Goodstein sequence starting in m is called the *Goodstein function*. That is, if n is the first index where $m_n = 0$, then $\mathcal{G}(m) = n + 1$ (since we index from zero).

Example. The first few values of the Goodstein function are

- $\mathcal{G}(1) = 2, \quad \mathcal{G}(2) = 4, \quad \mathcal{G}(3) = 6,$
- $\mathcal{G}(4) = 3 \cdot 2^{402\,653\,211} - 2 > 10^{10^8} = 10^{100\,000\,000}$,
- $\mathcal{G}(5) > 10^{10^{10^{10\,000}}}, \quad \mathcal{G}(12) >$ Graham's number.[13]

Observation 3.78. *The Goodstein function $\mathcal{G}$ dominates f_α for every $\alpha < \varepsilon_0$.*

Proof. Notice that we could reformulate Goodstein's theorem as follows: the function $\mathcal{G} : \omega \to \omega$ is total. Now, if $\mathcal{G}$ was dominated by some f_α, then PA could prove that $\mathcal{G}$ is total, thus proving Goodstein's theorem. □

Corollary 3.79. *The Goodstein function is not primitive recursive since it dominates the Ackerman function $A(n,n) \sim f_\omega(n)$.*

Theorem 3.80 (Cichon, 1983)**.** *The Goodstein function is dominated by H_{ε_0} (and thus also by f_{ε_0}). More precisely,*

$$\mathcal{G}(n) = H_{R_2^\omega(n+1)}(1) - 1,$$

where $R_2^\omega(n)$ is the result of writing n in hereditary base-2 notation and then replacing all 2s with ω (as we did when proving Goodstein's theorem).

[13] Numberphile has a VIDEO where Ron Graham himself explains the Graham's number.

It might not be immediately obvious that this is dominated by H_{ε_0}, but consider what happens when for example $n = 15 = 2^{2^2} - 1$. Then

$$\mathcal{G}(15) = H_{R_2^\omega(15+1)}(1) - 1 = H_{\omega^{\omega^\omega}}(1) - 1, \qquad H_{\varepsilon_0}(15) = H_{^{15}\omega}(15),$$

where $^{15}\omega$ denotes the ω-tower of height 15.

Theorem 3.81 (Caicedo, 2007)**.** *If $n = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_k}$ is the base-2 representation of n with $m_1 > m_2 > \cdots > m_k$, then*

$$\mathcal{G}(n) = f_{\alpha_1}\big(f_{\alpha_2}(\dots f_{\alpha_k}(3)\dots)\big) - 2,$$

where $\alpha_i = R_2^\omega(m_i)$.

4 Cardinal Numbers

Similar to how ordinal numbers represent order types of well-ordered sets, *cardinal numbers* represent *sizes* of well-ordered sets. Historically, these were two different classes of abstraction, but for us, cardinal numbers will be a special type of ordinal numbers.

4.1 Cardinals as Sizes of Well-Ordered Sets

We know that sizes of sets can be compared using injective mappings and that the relation $x \approx y$ is an equivalence on the class of all sets. We would now like to take the step from only *comparing* the sizes of sets to *quantifying* them. The question is, how do we represent classes

$$\{y \mid y \approx x\}$$

of all sets with the same size as x? We would like to define a mapping that assigns to each set x a set $|x|$ so that for any two sets x and y, we would have

$$x \approx y \iff |x| = |y|. \tag{4.1}$$

It would be ideal if, in addition, $x \approx |x|$. If such a mapping exists, then we call the set $|x|$ the *cardinality* of the set x.

It is fairly easy to define the cardinalities of certain classes of sets. For example, if x is finite, then there is a unique natural number n such that $x \approx n$, and we can set $|x| \coloneqq n$. Similarly, if x is countable, then $|x| \coloneqq \omega$. We are beginning to see a pattern: if x can be well-ordered, choose $|x|$ as the order type of one of its orderings; but which one? In order for (4.1) to hold, we have to choose $|x|$ as the *least* such order type. This is the key to defining cardinal numbers.

Definition 4.1 (Cardinal numbers)**.** An ordinal number κ is a *cardinal number* if for all ordinals $\alpha < \kappa$ there is no injection $\kappa \to \alpha$. Equivalently, if

$$\alpha < \kappa \implies \alpha \prec \kappa.$$

We denote the *class of all cardinal numbers* by Cn. ($\mathrm{Cn} \subset \mathrm{On}$).

We say that a cardinal number κ is the *cardinality* of a set x, and we write $|x| = \kappa$, if there exists a bijection $x \to \kappa$; in other words, if $x \approx \kappa$.

As for notation, we will usually denote cardinal numbers using letters from the middle of the Greek alphabet: $\kappa, \lambda, \mu, \nu \ldots$

Observation 4.2. *For cardinal numbers κ and λ it holds that*

(a) $\kappa < \lambda \iff \kappa \prec \lambda$,

(b) $\kappa = \lambda \iff \kappa \approx \lambda$,

Observation 4.3. *If the sets x and y have their cardinalities defined, then*

(a) $x \approx y \iff |x| = |y|$,

(b) $x \approx |x|$.

Observation 4.4. *The cardinality $|x|$ is defined $\iff$ x can be well-ordered.*

Proof. Our earlier discussion and the way we defined cardinals show that if x can be well-ordered, then $|x|$ is defined (as the least order type of the well-orderings of x). On the other hand, if $|x| = \kappa$ is defined, then we can well-order x by inheriting the order of κ. □

This demonstrates that if we want to have $|x|$ defined for every set x, then we need to accept the axiom of choice (which allows us to well-order any set).

Remark. There actually is a way to define a map assigning to every set x a set $|x|$ such that (a) holds *without* using the axiom of choice. This is achieved via the axiom of foundation,[14] but we lose property (b). A. Lévy (1969) showed that without either the axiom of foundation or the axiom of choice, it is impossible to define a map assigning each set x a set $|x|$ such that (a) holds. D. Pincus (1974) showed that without the axiom of choice, there exists no mapping satisfying both (a) and (b) for all sets x.

Example. Some basic properties of Cn are:

- every $n \in \omega$ and ω are cardinal numbers,
- if $\alpha \geq \omega$, then $\alpha + 1$ is not a cardinal number,
- thus every cardinal number $\kappa \geq \omega$ is a limit ordinal number,
- but not every limit ordinal number is a cardinal number — for example, we have $\omega + \omega > \omega$, but $\omega + \omega \approx \omega$, so $\omega + \omega$ is not a cardinal number,
- by the same logic, if $\alpha > \omega$ is countable, then it is not a cardinal number,
- the first uncountable ordinal ω_1 we encountered in Section 3.3.3 is a cardinal number since every $\alpha < \omega_1$ is countable.

Lemma 4.5. *If $A \subseteq \mathrm{Cn}$ is a set of cardinal numbers, then $\sup(A) = \bigcup A$ is also a cardinal number. In other words, the class $\mathrm{Cn} \subset \mathrm{On}$ is closed.*

[14] Briefly, this is done by showing that the axiom of foundation implies that the relation $\in$ is well-founded on V, allowing us to define a hierarchy of sets $V_0 \coloneqq \varnothing$, $V_{\alpha+1} \coloneqq \mathcal{P}(V_\alpha)$ and $V_\lambda \coloneqq \bigcup_{\alpha<\lambda} V_\alpha$ for limit λ, such that $V = \bigcup\{V_\alpha \mid \alpha \in \mathrm{On}\}$. Now, for any set x denote $\varrho(x) \coloneqq \min\{\alpha \mid x \subseteq V_\alpha\}$ and define $|x|$ as the set $\{y \mid y \approx x \wedge (\forall z)\big(z \approx x \Rightarrow \varrho(y) \leq \varrho(z)\big)\}$.

Proof. Since A is a set of ordinals, according to Lemma 3.4, $\sigma := \sup(A) = \bigcup A$ is an ordinal number. We need to show the implication

$$\alpha < \sigma \implies \alpha \prec \sigma.$$

Let $\alpha < \sigma$. Then $\alpha \in \sigma = \bigcup A$, so $\alpha \in \kappa$ for some cardinal $\kappa \in A$. Because κ is a cardinal number and $\alpha < \kappa$, it must be that $\alpha \prec \kappa$. Furthermore, because $\kappa \in A$, we have $\kappa \subseteq \bigcup A = \sigma$ and so $\kappa \preceq \sigma$. The Cantor–Bernstein theorem together with $\alpha \prec \kappa \preceq \sigma$ implies that $\alpha \prec \sigma$. □

Theorem 4.6. *For every cardinal there exists a larger cardinal.*

Proof. Suppose for contradiction that κ is the largest cardinal. Then, for every ordinal $\alpha \geq \kappa$, there exists a bijection $\alpha \to \kappa$, so κ can be well-ordered according to the type α. Corollary 2.12 implies that if r_α, r_β are well-orderings of κ with types $\alpha \neq \beta$, then $r_\alpha \neq r_\beta$ because α and β are not isomorphic.

Notice that every well-ordering of κ is a subset of $\kappa \times \kappa$, so an element of $\mathcal{P}(\kappa \times \kappa)$. For $\alpha \geq \kappa$, denote by $R_\alpha \in \mathcal{P}(\mathcal{P}(\kappa \times \kappa))$ the set of all well-orderings of κ with type α. Finally, notice that for $\alpha \neq \beta$ we have $R_\alpha \neq R_\beta$, so we can construct an injection $\alpha \mapsto R_\alpha$ that maps the proper class $\mathrm{On} \setminus \kappa$ into the set $\mathcal{P}(\mathcal{P}(\kappa \times \kappa))$, which contradicts the axiom schema of replacement. □

Corollary 4.7. *The class of all cardinal numbers* Cn *is a proper class.*

Proof. If it were a set, then by Lemma 4.5, its supremum would be the largest cardinal number, which is impossible. □

Definition 4.8. The *successor* of a cardinal κ is the smallest cardinal larger than κ, and we denote it by κ^+. Furthermore, we say that κ is the *predecessor* of κ^+. Finally, $\lambda > 0$ is a *limit* cardinal if it has no predecessor.

4.2 Infinite Cardinals

Finite cardinals are not very interesting, since they are the same as the natural numbers. We are interested in the cardinalities of infinite sets. The cardinals that measure the sizes of infinite (well-ordered) sets are the elements of the well-ordered proper class $\mathrm{Cn} \setminus \omega$, and by Exercise 5, there exists a unique increasing ordinal function $\aleph : \mathrm{On} \to \mathrm{Cn} \setminus \omega$ that is an isomorphism of the classes On and $\mathrm{Cn} \setminus \omega$. According to Lemma 4.5 it is continuous and therefore normal. Cantor introduced the symbol $\aleph$ ("aleph"), the first letter of the Hebrew alphabet, to denote this function. The smallest infinite cardinal number (the size of countable sets) is denoted as $\aleph_0$ and read "aleph null."

Definition 4.9. The unique normal function mapping On to the class of all infinite cardinals is denoted by $\aleph$, and its values $\aleph(\alpha)$ are denoted by $\aleph_\alpha$.

Observation 4.10. *These aleph numbers are exactly the omega numbers we have discovered in Section 3.3.3. That is,* $\aleph_\alpha = \omega_\alpha$ *and*

$$\aleph_\alpha = \{\beta \in \mathrm{On} \mid |\beta| < \aleph_\alpha\}$$

The cardinal $\aleph_\alpha$ *is the first ordinal number with cardinality* $\aleph_\alpha$.

Remark. Historically, ordinals and cardinals were not concrete sets but abstract concepts: ordinals described well-ordering types, while cardinals measured size. This distinction led to the development of two parallel notation systems, ω_α and $\aleph_\alpha$. Von Neumann's 1923 set-theoretic definition of ordinals unified these ideas by providing canonical representatives for ordinal types. Today, we often write ω_α when considering the cardinal $\aleph_\alpha$ viewed as an ordinal with its well-order.

Observation 4.11. *$\aleph_0 = \omega$ is a limit cardinal, and for $\alpha > 0$ we have that*

$$\aleph_\alpha \text{ is a limit cardinal} \iff \alpha \text{ is a limit ordinal.}$$

Proof. We use the fact that a cardinal is limit when it is not isolated and prove the statement by contraposition. The claim follows from the simple observation that when $\alpha = \beta + 1$ is an isolated ordinal, then $\aleph_\alpha = \aleph_{\beta+1} = \aleph_\beta^+$. And when $\aleph_\alpha = \aleph_\beta^+$ is an isolated cardinal, then $\alpha = \beta + 1$. □

Observation 4.12. *If α is an ordinal and ξ is a limit ordinal, then*

(a) $\alpha \leq \aleph_\alpha$, *... $\aleph$ is a normal function*

(b) *there exist ordinals α such that $\alpha = \aleph_\alpha$,* *... see Theorem 3.12*

(c) ω_α *is a limit ordinal,* *... it is an infinite cardinal*

(d) $\aleph_\xi = \sup\{\aleph_\alpha \mid \alpha < \xi\}$ *... $\aleph$ is a normal function*

Theorem 4.13. *For every ordinal α, it holds that $|\aleph_\alpha \times \aleph_\alpha| = \aleph_\alpha$.*

Corollary 4.14. *If x can be well-ordered, then $x \times x \approx x$. By induction, this also holds for any finite Cartesian product $x \times \cdots \times x \approx x$.*

To prove this theorem, we first define a suitable well-ordering of $\aleph_\alpha \times \aleph_\alpha$.

Definition 4.15. For ordinals α and β we define the *maximo-lexicographical* ordering of the set $\alpha \times \beta$ as

$$(\alpha_1,\beta_1) \sqsubset (\alpha_2,\beta_2) \iff \begin{cases} \max\{\alpha_1,\beta_1\} < \max\{\alpha_2,\beta_2\}, \text{ or} \\ \max\{\alpha_1,\beta_1\} = \max\{\alpha_2,\beta_2\} \wedge \alpha_1 < \alpha_2, \text{ or} \\ \max\{\alpha_1,\beta_1\} = \max\{\alpha_2,\beta_2\} \wedge \alpha_1 = \alpha_2 \wedge \beta_1 < \beta_2. \end{cases}$$

Intuition. Picture the product $\alpha \times \beta$ as a grid whose horizontal axis is indexed by α and whose vertical axis is indexed by β. Every point (α_i,β_j) sits in a "right-angle" band determined by the value $\max\{\alpha_i,\beta_j\}$. These bands consist of all points whose coordinates share the same maximum, and the bands themselves move outward from the origin. The ordering $\sqsubset$ simply compares points by the outward distance of the bands they belong to. Once two points lie in the same band, we break ties lexicographically. First, compare the α-coordinates. If those agree, compare the β-coordinates. Thus, inside each strip, the ordering "runs along the top" from left to right, and only then "climbs upward" on the right edge. It might be helpful to draw this on a piece of paper.

Exercise 16. Prove that the ordering defined above is a well-ordering.

Exercise 17. Prove that $\omega \times \omega \approx \omega$ by showing that $(\omega \times \omega, \sqsubset)$ and $(\omega, <)$ are order-isomorphic, using the result of Exercise 4. Although one can also show that $\omega \times \omega \approx \omega$ using prime-number encodings, this set-theoretic approach is conceptually cleaner, as it does not require arithmetic.

Proof of Theorem 4.13. By transfinite induction on α. If $\alpha = 0$, then we have the countable case $\omega \times \omega \approx \omega$, which holds by the previous exercise. Suppose that $\alpha > 0$ and consider the maximo-lexicographical ordering of $\omega_\alpha \times \omega_\alpha$. Since this is a well-ordering, it is isomorphic to a unique ordinal η, and we claim that $\eta = \omega_\alpha$. Clearly $\omega_\alpha \preceq \omega_\alpha \times \omega_\alpha \approx \eta$. This implies that $\omega_\alpha \leq \eta$ since ω_α is a cardinal number. Suppose for contradiction that $\omega_\alpha < \eta$; that is, $\omega_\alpha = (\leftarrow, \omega_\alpha) \subset \eta$ is isomorphic to an initial segment $(\leftarrow, (\gamma,\delta))$ of $(\omega_\alpha \times \omega_\alpha, \sqsubset)$, where $(\gamma, \delta) \in \omega_\alpha \times \omega_\alpha$.

Let $\xi = \max\{\gamma,\delta\}+1$ and notice that $(\leftarrow, (\gamma,\delta)) \subseteq \xi \times \xi$. Since ω_α is a cardinal and $\xi < \omega_\alpha$, there exists $\beta < \alpha$ such that $|\xi| = \omega_\beta$. By the induction hypothesis,

$$|\xi \times \xi| = |\omega_\beta \times \omega_\beta| = \omega_\beta < \omega_\alpha.$$

This is a contradiction since ω_α is isomorphic to an initial segment of $\xi \times \xi$, and we would have $\omega_\alpha < \omega_\alpha$. □

Theorem 4.16. AC $\iff$ *For every infinite set x, we have $x \times x \approx x$.*

Proof sketch. The direction '$\Rightarrow$' is easy; just well-order x and use Theorem 4.13.

The reverse implication is harder, and we only sketch the proof. We show that if A is an infinite set satisfying $A \times A \approx A$, then we can well-order A, implying the well-ordering principle. Let $j : A \times A \to A$ be a bijection, and for each $a \in A$ define $C_a \coloneqq \{j(a,t) \mid t \in A\}$. Notice that the family $\{C_a \mid a \in A\}$ partitions A into "A many" copies of A. In other words, the sets C_a are pairwise disjoint and satisfy $C_a \approx A$ and $\bigcup_{a\in A} C_a = A$.

Let η be the Hartogs number (see Theorem 3.44) of A; that is, the first ordinal that does not inject into A. The idea of the proof is to try to use transfinite recursion to build a family of well-ordered subsets $A_\alpha \subseteq A$ for each $\alpha < \eta$ such that the order type of A_α is α, and the sets A_α are pairwise disjoint. Informally, we do this by placing each A_α into the slot reserved by some C_a. Since $C_a \approx A$, the set A_α can always "fit" into C_a, and the sets C_a are pairwise disjoint. Suppose the transfinite recursion succeeds, and we define all pairwise disjoint A_α with isomorphisms $j_\alpha : \alpha \to A_\alpha$. Then we can define an injection $\eta \setminus \{\varnothing\} \to A$ as $\alpha \mapsto j_\alpha(0)$, contradicting the choice of η.

Hence, the transfinite recursion must "fail," and at some step $\alpha < \eta$, we will have already used up all slots C_a. That is, for every $\beta < \alpha$, there is a slot C_{a_β} containing A_β and $\bigcup_{\beta<\alpha} C_{a_\beta} = A$. Then we define a bijection $\alpha \to A$ as $\beta \mapsto a_\beta$, allowing us to well-order A. □

Theorem 4.17. *For any ordinal α it holds that*

(a) $\aleph_\alpha \prec \{x \mid x \subseteq \aleph_\alpha\} = \mathcal{P}(\aleph_\alpha)$,

(b) $\aleph_\alpha \approx \{x \mid x \subseteq \aleph_\alpha$ *is finite*$\} \subset \mathcal{P}(\aleph_\alpha)$.

Proof. The first claim follows from Cantor's theorem (the set of subsets of a set x cannot be injected into x). The second claim says that if we restrict ourselves to only finite subsets, then an injection into the original set is possible. We omit the proof, but it can be found in §4 Chapter II of [1]. □

Definition 4.18. Let κ and λ be cardinals. We define cardinal numbers

(a) $\kappa + \lambda \coloneqq |(\{0\} \times \kappa) \cup (\{1\} \times \lambda)|$,

(b) $\kappa \cdot \lambda = |\lambda \times \kappa|$.

In other words, $\kappa + \lambda$ and $\kappa \cdot \lambda$ are cardinal numbers, which represent the size of the set on the right side of the equation, in contrast to ordinal addition and multiplication, which express the order type of the same set when ordered lexicographically. If we want to highlight this difference, we talk about *cardinal* addition and multiplication.

Observation 4.19. *Cardinal addition and multiplication are associative, commutative, and distributive. Additionally, when restricted to ω, they are the same as the corresponding ordinal operations.*

Recall that ordinal addition and multiplication are associative; however, in general, they are not commutative or left-distributive. This is because they have to "keep track" of the underlying orderings.

Lemma 4.20. *If κ and λ are cardinals, and at least one of them is infinite, then $\kappa + \lambda = \max\{\kappa,\lambda\}$. If, in addition, they are nonzero, then $\kappa \cdot \lambda = \max\{\kappa,\lambda\}$.*

Proof. Denote $\mu \coloneqq \max\{\kappa,\lambda\}$. To show addition, consider

$$\begin{aligned}\mu \preceq \kappa + \lambda &\approx (\{0\} \times \kappa) \cup (\{1\} \times \lambda) \\ &\preceq (\{0\} \times \mu) \cup (\{1\} \times \mu) = 2 \times \mu \preceq \mu \times \mu \approx \mu.\end{aligned}$$

where the last '$\approx$' follows from Theorem 4.13. The Cantor–Bernstein theorem now implies that $\kappa + \lambda \approx \mu$, so $\kappa + \lambda = \mu$ (because they are cardinals).

Similarly for multiplication: $\mu \preceq \kappa \cdot \lambda \approx \lambda \times \kappa \preceq \mu \times \mu \approx \mu$. □

Corollary 4.21. *For any ordinals α and β, it holds that*

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \max\{\aleph_\alpha,\aleph_\beta\}.$$

Lemma 4.22 (AC)**.** *For any set S, we have $|\bigcup S| \leq |S| \cdot \sup\{|A| \mid A \in S\}$.*

Proof. Let $\kappa = |S|$ and $\lambda = \sup\{|A| \mid A \in S\}$. Since $\kappa \approx S$, we can enumerate the elements of S as $S = \{A_\alpha \mid \alpha < \kappa\}$. Moreover, each A_α injects into λ; hence, we can *choose* an injection $j_\alpha : A_\alpha \to \lambda$. For an element $a \in \bigcup S$, define

$$\alpha_a \coloneqq \min\{\alpha < \kappa \mid a \in A_\alpha\}.$$

This number indicates in which A_α does a first appear in. Notice that more elements $a \in \bigcup S$ can have the same number α_a, but that $j_{\alpha_a}(a)$ uniquely identifies a among these elements (since j_{α_a} is injective). This allows us to define an injection $g : \bigcup S \to \kappa \times \lambda$ as $a \mapsto (\alpha_a, j_{\alpha_a}(a))$. □

Corollary 4.23 (AC)**.** *The union of a collection of $\aleph_\alpha$ sets, each of cardinality at most $\aleph_\alpha$, has cardinality at most $\aleph_\alpha$. If, in addition, they are non-empty, then the union has cardinality exactly $\aleph_\alpha$.*

4.3 Cofinality and Inaccessible Cardinals

Recall the pigeonhole principle for $\omega = \aleph_0$, which says that ω cannot be partitioned into a finite number of finite sets. Thus, if $A = \bigcup\{A_i \mid i \in I\}$ is countably infinite, then either I or one of A_i has to be countably infinite. How does this generalize to higher cardinals?

Definition 4.24. Let $(X, \leq_R)$ be a partially ordered set. We say that $Y \subseteq X$ is *cofinal* in X (or is a *cofinal subset* of X) with respect to $\leq_R$ if every $x \in X$ is bounded by some $y \in Y$; (that is, $x \leq y$).

Observation 4.25. *If Y is cofinal in X, then it contains all maximal elements of X. Moreover, the relation "to be cofinal in" is transitive.*

Example. If X has a maximum x, then $\{x\}$ is the smallest cofinal subset of X.

We are usually interested in cofinality in the context of limit ordinals.

Observation 4.26. *If α is an ordinal, $A \subseteq \alpha$, and*

(i) *$\alpha = \beta + 1$ is isolated, then A is cofinal in $\alpha \iff \beta \in A$,*

(ii) *α is limit, then A is cofinal in $\alpha \iff \sup(A) = \alpha$.*

Definition 4.27 (Cofinality)**.** The *cofinality* of a limit ordinal α, denoted by $\mathrm{cf}(\alpha)$, is the least ordinal β that is the order type of some $A \subseteq \alpha$ cofinal with α.

Lemma 4.28. *Cofinality of α is the length of the shortest sequence with limit α:*

$$\mathrm{cf}(\alpha) = \min\{|A| \mid A \subseteq \alpha \wedge \sup(A) = \alpha\}.$$

More precisely, $\mathrm{cf}(\alpha)$ is always an infinite cardinal $\aleph_\beta$ for some $\beta \geq 0$.

Proof. We know that $\mathrm{cf}(\alpha)$ is an infinite ordinal. If $\mathrm{cf}(\alpha)$ were not a cardinal number, then there would be a cardinal $\kappa < \mathrm{cf}(\alpha)$ such that $\kappa \approx \mathrm{cf}(\alpha)$, and hence a bijection $f : \kappa \to \mathrm{cf}(\alpha)$. Define a function $g : \kappa \to \mathrm{cf}(\alpha)$ as $g : \beta \mapsto \sup\{f(\alpha) \mid \alpha < \beta\}$, and notice that g is non-decreasing and $\mathrm{Rng}(g)$ is a cofinal subset of $\mathrm{cf}(\alpha)$ ordered according to some type $\gamma \leq \kappa$.

Since $\mathrm{cf}(\alpha)$ is the cofinality of α, there exists a subset $A \subseteq \alpha$ ordered according to $\mathrm{cf}(\alpha)$, and an isomorphism $h : \mathrm{cf}(\alpha) \to A$. Because h preserves order, it is increasing; therefore, $h[\mathrm{Rng}(g)] \subseteq A$ is a cofinal subset of α, ordered by the type $\gamma \leq \kappa < \mathrm{cf}(\alpha)$, which is a contradiction. □

Corollary 4.29. *For any (limit) ordinal α it holds that $\omega \leq \mathrm{cf}(\alpha) \leq |\alpha|$.*

Example. Some cofinalities we already know:

- $\mathrm{cf}(\omega) = \mathrm{cf}(\omega + \omega) = \mathrm{cf}(\omega \cdot \omega) = \mathrm{cf}(\omega^\omega) = \mathrm{cf}(\varepsilon_0) = \mathrm{cf}(\Gamma_0) = \omega$,
- in general, $\mathrm{cf}(\alpha) = \omega$ for countable (limit) α, because $\omega \leq \mathrm{cf}(\alpha)$, and $\sup(\alpha) = \alpha$, so $\mathrm{cf}(\alpha) \leq |\alpha| = \omega$,
- $\mathrm{cf}(\aleph_\omega) = \omega$, since $\aleph_\omega = \sup\{\aleph_\alpha \mid \alpha < \omega\}$,
- $\mathsf{AC}_\omega \implies \mathrm{cf}(\omega_1) = \omega_1$, as Lemma 3.35 implies $\mathrm{cf}\,\omega_1 \geq \omega_1$.

Lemma 4.30. *For every (limit) ordinal α we have*

$$\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha).$$

Proof. Let $\beta = \mathrm{cf}(\alpha)$ and $\gamma = \mathrm{cf}(\mathrm{cf}(\alpha))$, so $\gamma \leq \beta$. From the definition of $\mathrm{cf}(\alpha)$, there exists a cofinal subset $A \subseteq \alpha$ and an increasing function (isomorphism) $f : \beta \to A$, and a cofinal subset $B \subseteq \beta$ and an increasing function (isomorphism) $g : \gamma \to B$. Notice that the map $g \circ f : \gamma \to A$ is an isomorphism of γ and a cofinal subset of α, thus (by the definition of $\mathrm{cf}(\alpha)$), $\gamma \geq \mathrm{cf}(\alpha) = \beta$. □

Definition 4.31. An infinite cardinal number $\kappa = \aleph_\alpha$ is a

(a) *regular cardinal* if $\mathrm{cf}(\kappa) = \kappa$,

(b) *singular cardinal* if $\mathrm{cf}(\kappa) < \kappa$.

Intuition. If ω_α is a regular cardinal, then it is *almost* closed on taking suprema. As long as the length of the sequence is less than $\aleph_\alpha$, the limit will never reach ω_α. However, if ω_α is singular, it is possible to reach it from bellow via a shorter sequence.

We have already seen that $\aleph_0 = \omega$ is regular and $\aleph_\omega$ singular. Furthermore, Lemma 4.30 implies that the cofinality of any (limit) ordinal α is always a regular cardinal number. The question is: are there any regular cardinals besides $\aleph_0$? If we assume the axiom of countable choice, then ω_1 is also regular. But what if we are working in bare ZF?

Gitik (1979) showed that the statement "$\aleph_0$ is the only regular cardinal" cannot be disproved in ZF. In that case, *every* limit ordinal α has a cofinal subset of size $\aleph_0$, and there exists a sequence $(\gamma_n)_{n<\omega}$ with limit α.

Theorem 4.32. *An infinite cardinal number κ is singular $\iff$ there exists a set X such that $\kappa = \bigcup X$ where $|X| < \kappa$ and $(\forall x \in X)(|x| < \kappa)$.*

Proof. TODO □

Corollary 4.33 (Pigeonhole principle for cardinals)**.** *If κ is a regular cardinal and $\kappa = \bigcup X$, where $|X| < \kappa$, then there exists $x \in X$ such that $|x| = \kappa$.*

If we further assume the axiom of choice, every set will have a defined cardinality, and we can make a more general claim.

Corollary 4.34 (Pigeonhole principle for infinite sets; AC)**.** *If S is an infinite set with regular cardinality $|S|$, and $S = \bigcup X$, where $|X| < |S|$, then there exists $x \in X$ such that $|x| = |S|$.*

Theorem 4.35 (AC)**.** *Every infinite isolated cardinal $\aleph_{\alpha+1}$ is regular.*

Proof. If it were singular, then by Theorem 4.32 it would be the supremum of at most $\aleph_\alpha$ sets of cardinality at most $\aleph_\alpha$. But Corollary 4.23 implies that such a supremum must have cardinality at most $\aleph_\alpha$ (since suprema and unions of ordinals are the same). □

Theorem 4.36. *If $\aleph_\alpha > \aleph_0$ is a limit cardinal (α a limit ordinal), then*

$$\mathrm{cf}(\aleph_\alpha) = \mathrm{cf}(\alpha).$$

Proof. The cardinal $\aleph_\alpha$ is defined as the limit of the sequence $\{\aleph_\beta \mid \beta < \alpha\}$. The claim follows from the observation that the set $\{\aleph_\beta \mid \beta \in A\}$ is cofinal in $\aleph_\alpha$ if and only if A is a cofinal subset of α. This allows us to skip some terms of the sequence; in fact, we only need $\mathrm{cf}(\alpha)$ many terms. □

We know that in ZFC, every infinite isolated cardinal is regular, but we have not seen any regular limit cardinals besides $\aleph_0$. Let's try different limit ordinal indices α and check whether $\aleph_\alpha$ is regular or not.

- $\mathrm{cf}(\aleph_\omega) = \mathrm{cf}(\aleph_{\varepsilon_0}) = \mathrm{cf}(\aleph_{\Gamma_0}) = \omega$,
- $\mathrm{cf}(\aleph_\alpha) = \omega$ for any countable (limit) α,
- $\mathrm{cf}(\aleph_{\omega_1}) = \mathrm{cf}(\omega_1) = \aleph_1$,
- $\mathrm{cf}(\aleph_{\omega_{\alpha+1}}) = \mathrm{cf}(\omega_{\alpha+1}) = \aleph_{\alpha+1}$ for any isolated cardinal $\omega_{\alpha+1}$.

This clearly is not working. Notice that Theorem 4.36 implies that if $\aleph_\alpha > \aleph_0$ is a regular limit cardinal, then

$$\mathrm{cf}(\aleph_\alpha) = \mathrm{cf}(\alpha) = \aleph_\alpha.$$

Since $\mathrm{cf}(\alpha) \leq \alpha$ and $\aleph_\alpha \geq \alpha$, we conclude that $\alpha = \aleph_\alpha$. Hence, every regular limit cardinal larger than $\aleph_0$ has to be a fixed point of the aleph function.

Theorem 3.12 allows us to construct fixed points of $\aleph$ quite easily; for example, the first fixed point is the limit of the sequence

$$\kappa_0 = 0,\ \kappa_{n+1} = \aleph_{\kappa_n} \qquad \longrightarrow \qquad 0, \aleph_0, \aleph_{\aleph_0}, \aleph_{\aleph_{\aleph_0}}, \aleph_{\aleph_{\aleph_{\aleph_0}}}, \dots$$

Let us denote it κ_0 and notice that it is singular since $\mathrm{cf}(\kappa_0) = \omega$. Similarly, any fixed point constructed using this method will have cofinality ω. The question is: can a fixed point of $\aleph$ be a regular cardinal? People started using the term *inaccessible* to describe these cardinals: they cannot be reached from below either by taking successor cardinals or by forming suprema of smaller cardinals. Today, we call such numbers *weakly inaccessible*.

Definition 4.37. A cardinal κ is *weakly inaccessible* if it is uncountable, regular, and a limit cardinal.

Remark. Weakly inaccessible cardinals were first introduced by F. Hausdorff in 1908. Only much later was it shown that the existence of such cardinals cannot be proved in ZFC (provided it is consistent). Therefore, it is consistent with ZFC to assume that every uncountable limit cardinal number is singular.

There are many problems in set theory which lead to the question if there exist large enough cardinal numbers to satisfy certain properties. If the existence of such numbers cannot be proven (for example, because such a number would be larger than a weakly inaccessible cardinal), we say that it is a *large cardinal*. There is an entire hierarchy of large cardinals, and the weakly inaccessible cardinal is the smallest of them all.

Theorem 4.38. ZFC *cannot prove the existence of weakly inaccessible cardinals.*

Proof sketch. This is a consequence of Gödel's second incompleteness theorem. At the end of Section 3.4.3 we mentioned that Gödel introduced the *constructible universe* L, which is a proper class built in stages. Informally, denote by $\mathrm{Df}(X)$ the class of sets definable by formulas with parameters from X. Then define

$$L_0 \coloneqq \varnothing, \quad L_{\alpha+1} \coloneqq \mathrm{Df}(L_\alpha \cup \{L_\alpha\}) \cap \mathcal{P}(L_\alpha), \quad L_\lambda \coloneqq \bigcup \{L_\alpha \mid \alpha < \lambda\},$$

where λ is a limit ordinal. We can check from outside of ZFC that the class $L = \bigcup L_\alpha$ satisfies all axioms and is a model. However, if κ is a weakly inaccessible cardinal, then the set L_κ is already "closed enough" to satisfy the axioms. Now, if ZFC could prove that κ exists, it could prove that it has a model (L_κ), thus proving its own consistency, contradicting Gödel's second incompleteness theorem. □

4.4 The Continuum Hypothesis

4.5 Cardinal Arithmetic

Sources

This document serves as lecture notes for the course NMAI074 taught at MFF CUNI by doc. Kynčl. The web of the course is HERE. A significant portion of the text follows parts of the second and third chapters of [1], which is in Czech. My notes from the introductory set theory course can be found HERE, also in Czech.

If you found any mistakes or errors, please contact me at smolikj@matfyz.cz.

[1] Petr Balcar Bohuslav a Štěpánek. *Teorie množin*. Vydání 2., opravené a rozšířené. Praha: Academia, 2001. ISBN: 80-200-0470-X.

[2] Stephen Budiansky. *Journey to the Edge of Reason: The Life of Kurt Gödel*. W. W. Norton, 2021. ISBN: 9781324005452.

[3] Jakub Bulín. *NAIL062 Propositional and Predicate Logic: Lecture Notes*. 2025, pp. 149–164. URL: https://github.com/jbulin-mff-uk/nail062/raw/main/lecture/lecture-notes/lecture-notes.pdf (visited on 11/06/2025).

[4] Jochen Burghardt. *Matchstick representation of ordinal numbers up to ω^ω*. 2023. URL: https://commons.wikimedia.org/wiki/File:Omega-exp-omega-normal.pdf (visited on 11/04/2025).

[5] Jochen Burghardt. *Order Type Examples*. 2019. URL: https://commons.wikimedia.org/wiki/File:OrderTypeExamples.pdf (visited on 11/04/2025).

[6] Pop-up casket. *A spiral representation of ordinals less than ω^ω*. 2012. URL: https://en.wikipedia.org/wiki/File:Omega-exp-omega-labeled.svg (visited on 11/04/2025).

[7] Timothy Y Chow. "The consistency of arithmetic". In: *The Mathematical Intelligencer* 41.1 (2019), pp. 22–30. URL: https://arxiv.org/pdf/1807.05641.

[8] Walter Dean and Sean Walsh. "The prehistory of the subsystems of second-order arithmetic". In: *The Review of Symbolic Logic* 10.2 (2017), pp. 357–396. URL: https://arxiv.org/pdf/1612.06219.

[9] Herbert B Enderton. *Elements of set theory*. Gulf Professional Publishing, 1977, pp. 195–196.

[10] Jean H Gallier. "What's so special about Kruskal's theorem and the ordinal Γ_0? A survey of some results in proof theory". In: *Annals of pure and applied logic* 53.3 (1991), pp. 199–260. URL: https://www.cis.upenn.edu/~jean/kruskal.pdf.

[11] Gina Garcia Tarrach. "The Axiom of Choice and its implications in mathematics". MA thesis. Universitat de Barcelona, 2017. URL: https://hdl.handle.net/2445/121981.

[12] Kurt Gödel. "On Formally Undecidable Propositions of Principia Mathematica and Related Systems I". In: *Kurt Gödel: Collected Works: Volume I Publications 1929-1936*. OUP Oxford, 1986, pp. 144–195. ISBN: 0-19-503964-5.

[13] Reuben Louis Goodstein. "On the restricted ordinal theorem". In: *The Journal of Symbolic Logic* 9.2 (1944), pp. 33–41. DOI: 10.2307/2268019.

[14] Gro-Tsen and IkamusumeFan. *A graphical "matchstick" representation of the ordinal ω^2*. 2015. URL: https://commons.wikimedia.org/wiki/File:Ordinal_ww.svg (visited on 11/04/2025).

[15] Joel David Hamkins. *Transfinite recursion as a fundamental principle in set theory*. Accessed: 2025-11-08. 2014. URL: https://jdh.hamkins.org/transfinite-recursion-as-a-fundamental-principle-in-set-theory/.

[16] Karel Hrbáček and Tomáš Jech. *Introduction to set theory*. eng. Third edition, revised and expanded. Pure and applied mathematics. A series of monographs and textbooks ; 220. Boca Raton: Taylor & Francis, 1999. ISBN: 0-8247-7915-0.

[17] Laurie Kirby and Jeff Paris. "Accessible independence results for Peano arithmetic". In: *Bulletin of the London Mathematical Society* 14.4 (1982), pp. 285–293. DOI: 10.1112/blms/14.4.285.

[18] Alberto Marcone and Antonio Montalbán. "The Veblen functions for computability theorists". In: *The Journal of symbolic logic* 76.2 (2011), pp. 575–602. URL: https://arxiv.org/pdf/0910.5442.

[19] Jeff Paris and Leo Harrington. "A mathematical incompleteness in Peano arithmetic". In: *Handbook of Mathematical Logic, edited by J. Barwise*. North-Holland, 1977, pp. 1133–1142.

[20] Michael Rathjen. "The art of ordinal analysis". In: *Proceedings of the International Congress of Mathematicians*. Vol. 2. European Mathematical Society. 2006, pp. 45–69.

[21] Stephen George Simpson. *Subsystems of second order arithmetic*. Vol. 1. Cambridge University Press, 2009.

[22] Will Sladek. *The termite and the tower: Goodstein sequences and provability in PA*. 2007. URL: https://andrescaicedo.wordpress.com/wp-content/uploads/2017/09/sladekgoodstein.pdf.

[23] Alan Mathison Turing et al. "On computable numbers, with an application to the Entscheidungsproblem". In: *J. of Math* 58.345-363 (1936), p. 5.

[24] Veritasium. *Math's Fundamental Flaw*. 2021. URL: https://www.youtube.com/watch?v=HeQX2HjkcNo (visited on 11/06/2025).

[25] Veritasium. *The Man Who Almost Broke Math (And Himself...) - Axiom of Choice*. 2025. URL: https://www.youtube.com/watch?v=_cr46G2K5Fo (visited on 11/08/2025).

[26] Vsauce. *How To Count Past Infinity*. 2016. URL: https://www.youtube.com/watch?v=SrU9YDoXE88 (visited on 11/04/2025).

[27] Googology Wiki. *Veblen function*. 2025. URL: https://googology.fandom.com/wiki/Veblen_function (visited on 11/12/2025).

[28] Wikipedia. *Hausdorff maximal principle — Wikipedia, The Free Encyclopedia*. 2025. URL: https://en.wikipedia.org/w/index.php?title=Hausdorff%5C%20maximal%5C%20principle&oldid=1300391387 (visited on 10/07/2025).